UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1996

GRADUATE STUDENT RESEARCH PROGRAM FINAL REPORTS

VOLUME 7B

ARMSTRONG LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES
5800 Uplander Way
Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Major Linda Steel-Goodwin

Program Manager, RDL
Scott Licoscos

Program Administrator, RDL
Johnetta Thompson

Program Administrator, RDL
Rebecca Kelly

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Bolling Air Force Base
Washington, D.C.
December 1996

REPORT DOCUMENTATION PAGE

Form Approved

AFRL-SR-BL-TR-00-

0737

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
December, 1996

3. REPO[…]

4. TITLE AND SUBTITLE
1996 Summer Research Program (SRP), Graduate Student Research Program (GSRP), Final Reports, Volume 7B, Armstrong Laboratory

5. FUNDING NUMBERS
F49620-93-C-0063

6. AUTHOR(S)
Gary Moore

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Research & Development Laboratories (RDL)
5800 Uplander Way
Culver City, CA 90230-6608

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research (AFOSR)
801 N. Randolph St.
Arlington, VA 22203-1977

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
The United States Air Force Summer Research Program (USAF-SRP) is designed to introduce university, college, and technical institute faculty members, graduate students, and high school students to Air Force research. This is accomplished by the faculty members (Summer Faculty Research Program, (SFRP)), graduate students (Graduate Student Research Program (GSRP)), and high school students (High School Apprenticeship Program (HSAP)) being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory (AFRL) Technical Directorates, Air Force Air Logistics Centers (ALC), and other AF Laboratories. This volume consists of a program overview, program management statistics, and the final technical reports from the GSRP participants at the Armstrong Laboratory.

14. SUBJECT TERMS
Air Force Research, Air Force, Engineering, Laboratories, Reports, Summer, Universities, Faculty, Graduate Student, High School Student

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

Due to its length, Volume 7 is bound in three parts, 7A, and 7B. Volume 7A contains #1-23. Volume 7B contains reports #24-35. The Table of Contents for Volume 7 is included in all parts.

This document is one of a set of 16 volumes describing the 1996 AFOSR Summer Research Program. The following volumes comprise the set:

VOLUME	**TITLE**
1	Program Management Report
	Summer Faculty Research Program (SFRP) Reports
2A & 2B	Armstrong Laboratory
3A & 3B	Phillips Laboratory
4	Rome Laboratory
5A , 5B & 5C	Wright Laboratory
6	Arnold Engineering Development Center, Wilford Hall Medical Center and Air Logistics Centers
	Graduate Student Research Program (GSRP) Reports
7A & 7B	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10A & 10B	Wright Laboratory
11	Arnold Engineering Development Center, United States Air Force Academy, Wilford Hall Medical Center, and Wright Patterson Medical Center
	High School Apprenticeship Program (HSAP) Reports
12A & 12B	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A&15B	Wright Laboratory
16	Arnold Engineering Development Center

GSRP FINAL REPORT TABLE OF CONTENTS i-x

<u>APPENDICIES:</u>

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

INTRODUCTION

The Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), offers paid opportunities for university faculty, graduate students, and high school students to conduct research in U.S. Air Force research laboratories nationwide during the summer.

Introduced by AFOSR in 1978, this innovative program is based on the concept of teaming academic researchers with Air Force scientists in the same disciplines using laboratory facilities and equipment not often available at associates' institutions.

The Summer Faculty Research Program (SFRP) is open annually to approximately 150 faculty members with at least two years of teaching and/or research experience in accredited U.S. colleges, universities, or technical institutions. SFRP associates must be either U.S. citizens or permanent residents.

The Graduate Student Research Program (GSRP) is open annually to approximately 100 graduate students holding a bachelor's or a master's degree; GSRP associates must be U.S. citizens enrolled full time at an accredited institution.

The High School Apprentice Program (HSAP) annually selects about 125 high school students located within a twenty mile commuting distance of participating Air Force laboratories.

AFOSR also offers its research associates an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. In 1994 the maximum amount of each grant was increased from $20,000 to $25,000, and the number of AFOSR-sponsored grants decreased from 75 to 60. A separate annual report is compiled on the SREP.

The numbers of projected summer research participants in each of the three categories and SREP "grants" are usually increased through direct sponsorship by participating laboratories.

AFOSR's SRP has well served its objectives of building critical links between Air Force research laboratories and the academic community, opening avenues of communications and forging new research relationships between Air Force and academic technical experts in areas of national interest, and strengthening the nation's efforts to sustain careers in science and engineering. The success of the SRP can be gauged from its growth from inception (see Table 1) and from the favorable responses the 1996 participants expressed in end-of-tour SRP evaluations (Appendix B).

AFOSR contracts for administration of the SRP by civilian contractors. The contract was first awarded to Research & Development Laboratories (RDL) in September 1990. After

completion of the 1990 contract, RDL (in 1993) won the recompetition for the basic year and four 1-year options.

2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM

The SRP began with faculty associates in 1979; graduate students were added in 1982 and high school students in 1986. The following table shows the number of associates in the program each year.

YEAR	SRP Participation, by Year			TOTAL
	SFRP	GSRP	HSAP	
1979	70			70
1980	87			87
1981	87			87
1982	91	17		108
1983	101	53		154
1984	152	84		236
1985	154	92		246
1986	158	100	42	300
1987	159	101	73	333
1988	153	107	101	361
1989	168	102	103	373
1990	165	121	132	418
1991	170	142	132	444
1992	185	121	159	464
1993	187	117	136	440
1994	192	117	133	442
1995	190	115	137	442
1996	188	109	138	435

Beginning in 1993, due to budget cuts, some of the laboratories weren't able to afford to fund as many associates as in previous years. Since then, the number of funded positions has remained fairly constant at a slightly lower level.

3. RECRUITING AND SELECTION

The SRP is conducted on a nationally advertised and competitive-selection basis. The advertising for faculty and graduate students consisted primarily of the mailing of 8,000 52-page SRP brochures to chairpersons of departments relevant to AFOSR research and to administrators of grants in accredited universities, colleges, and technical institutions. Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) were included. Brochures also went to all participating USAF laboratories, the previous year's participants, and numerous individual requesters (over 1000 annually).

RDL placed advertisements in the following publications: *Black Issues in Higher Education, Winds of Change,* and *IEEE Spectrum.* Because no participants list either *Physics Today* or *Chemical & Engineering News* as being their source of learning about the program for the past several years, advertisements in these magazines were dropped, and the funds were used to cover increases in brochure printing costs.

High school applicants can participate only in laboratories located no more than 20 miles from their residence. Tailored brochures on the HSAP were sent to the head counselors of 180 high schools in the vicinity of participating laboratories, with instructions for publicizing the program in their schools. High school students selected to serve at Wright Laboratory's Armament Directorate (Eglin Air Force Base, Florida) serve eleven weeks as opposed to the eight weeks normally worked by high school students at all other participating laboratories.

Each SFRP or GSRP applicant is given a first, second, and third choice of laboratory. High school students who have more than one laboratory or directorate near their homes are also given first, second, and third choices.

Laboratories make their selections and prioritize their nominees. AFOSR then determines the number to be funded at each laboratory and approves laboratories' selections.

Subsequently, laboratories use their own funds to sponsor additional candidates. Some selectees do not accept the appointment, so alternate candidates are chosen. This multi-step selection procedure results in some candidates being notified of their acceptance after scheduled deadlines. The total applicants and participants for 1996 are shown in this table.

1996 Applicants and Participants			
PARTICIPANT CATEGORY	TOTAL APPLICANTS	SELECTEES	DECLINING SELECTEES
SFRP	572	188	39
(HBCU/MI)	(119)	(27)	(5)
GSRP	235	109	7
(HBCU/MI)	(18)	(7)	(1)
HSAP	474	138	8
TOTAL	1281	435	54

4. SITE VISITS

During June and July of 1996, representatives of both AFOSR/NI and RDL visited each participating laboratory to provide briefings, answer questions, and resolve problems for both laboratory personnel and participants. The objective was to ensure that the SRP would be as constructive as possible for all participants. Both SRP participants and RDL representatives found these visits beneficial. At many of the laboratories, this was the only opportunity for all participants to meet at one time to share their experiences and exchange ideas.

5. HISTORICALLY BLACK COLLEGES AND UNIVERSITIES AND MINORITY INSTITUTIONS (HBCU/MIs)

Before 1993, an RDL program representative visited from seven to ten different HBCU/Mis annually to promote interest in the SRP among the faculty and graduate students. These efforts were marginally effective, yielding a doubling of HBCI/MI applicants. In an effort to achieve AFOSR's goal of 10% of all applicants and selectees being HBCU/MI qualified, the RDL team decided to try other avenues of approach to increase the number of qualified applicants. Through the combined efforts of the AFOSR Program Office at Bolling AFB and RDL, two very active minority groups were found, HACU (Hispanic American Colleges and Universities) and AISES (American Indian Science and Engineering Society). RDL is in communication with representatives of each of these organizations on a monthly basis to keep up with the their activities and special events. Both organizations have widely-distributed magazines/quarterlies in which RDL placed ads.

Since 1994 the number of both SFRP and GSRP HBCU/MI applicants and participants has increased ten-fold, from about two dozen SFRP applicants and a half dozen selectees to over 100 applicants and two dozen selectees, and a half-dozen GSRP applicants and two or three selectees to 18 applicants and 7 or 8 selectees. Since 1993, the SFRP had a two-fold applicant

increase and a two-fold selectee increase. Since 1993, the GSRP had a three-fold applicant increase and a three to four-fold increase in selectees.

In addition to RDL's special recruiting efforts, AFOSR attempts each year to obtain additional funding or use leftover funding from cancellations the past year to fund HBCU/MI associates. This year, 5 HBCU/MI SFRPs declined after they were selected (and there was no one qualified to replace them with). The following table records HBCU/MI participation in this program.

SRP HBCU/MI Participation, By Year				
YEAR	SFRP		GSRP	
	Applicants	Participants	Applicants	Participants
1985	76	23	15	11
1986	70	18	20	10
1987	82	32	32	10
1988	53	17	23	14
1989	39	15	13	4
1990	43	14	17	3
1991	42	13	8	5
1992	70	13	9	5
1993	60	13	6	2
1994	90	16	11	6
1995	90	21	20	8
1996	119	27	18	7

6. SRP FUNDING SOURCES

Funding sources for the 1996 SRP were the AFOSR-provided slots for the basic contract and laboratory funds. Funding sources by category for the 1996 SRP selected participants are shown here.

1996 SRP FUNDING CATEGORY	SFRP	GSRP	HSAP
AFOSR Basic Allocation Funds	141	85	123
USAF Laboratory Funds	37	19	15
HBCU/MI By AFOSR (Using Procured Addn'l Funds)	10	5	0
TOTAL	188	109	138

SFRP - 150 were selected, but nine canceled too late to be replaced.
GSRP - 90 were selected, but five canceled too late to be replaced (10 allocations for the ALCs were withheld by AFOSR.)
HSAP - 125 were selected, but two canceled too late to be replaced.

7. COMPENSATION FOR PARTICIPANTS

Compensation for SRP participants, per five-day work week, is shown in this table.

1996 SRP Associate Compensation

PARTICIPANT CATEGORY	1991	1992	1993	1994	1995	1996
Faculty Members	$690	$718	$740	$740	$740	$770
Graduate Student (Master's Degree)	$425	$442	$455	$455	$455	$470
Graduate Student (Bachelor's Degree)	$365	$380	$391	$391	$391	$400
High School Student (First Year)	$200	$200	$200	$200	$200	$200
High School Student (Subsequent Years)	$240	$240	$240	$240	$240	$240

The program also offered associates whose homes were more than 50 miles from the laboratory an expense allowance (seven days per week) of $50/day for faculty and $40/day for graduate students. Transportation to the laboratory at the beginning of their tour and back to their home destinations at the end was also reimbursed for these participants. Of the combined SFRP and

GSRP associates, 65 % (194 out of 297) claimed travel reimbursements at an average round-trip cost of $780.

Faculty members were encouraged to visit their laboratories before their summer tour began. All costs of these orientation visits were reimbursed. Forty-five percent (85 out of 188) of faculty associates took orientation trips at an average cost of $444. By contrast, in 1993, 58 % of SFRP associates took orientation visits at an average cost of $685; that was the highest percentage of associates opting to take an orientation trip since RDL has administered the SRP, and the highest average cost of an orientation trip. These 1993 numbers are included to show the fluctuation which can occur in these numbers for planning purposes.

Program participants submitted biweekly vouchers countersigned by their laboratory research focal point, and RDL issued paychecks so as to arrive in associates' hands two weeks later.

In 1996, RDL implemented direct deposit as a payment option for SFRP and GSRP associates. There were some growing pains. Of the 128 associates who opted for direct deposit, 17 did not check to ensure that their financial institutions could support direct deposit (and they couldn't), and eight associates never did provide RDL with their banks' ABA number (direct deposit bank routing number), so only 103 associates actually participated in the direct deposit program. The remaining associates received their stipend and expense payments via checks sent in the US mail.

HSAP program participants were considered actual RDL employees, and their respective state and federal income tax and Social Security were withheld from their paychecks. By the nature of their independent research, SFRP and GSRP program participants were considered to be consultants or independent contractors. As such, SFRP and GSRP associates were responsible for their own income taxes, Social Security, and insurance.

8. CONTENTS OF THE 1996 REPORT

The complete set of reports for the 1996 SRP includes this program management report (Volume 1) augmented by fifteen volumes of final research reports by the 1996 associates, as indicated below:

1996 SRP Final Report Volume Assignments

LABORATORY	SFRP	GSRP	HSAP
Armstrong	2	7	12
Phillips	3	8	13
Rome	4	9	14
Wright	5A, 5B	10	15
AEDC, ALCs, WHMC	6	11	16

APPENDIX A – PROGRAM STATISTICAL SUMMARY

A. Colleges/Universities Represented

Selected SFRP associates represented 169 different colleges, universities, and institutions, GSRP associates represented 95 different colleges, universities, and institutions.

B. States Represented

SFRP -Applicants came from 47 states plus Washington D.C. and Puerto Rico. Selectees represent 44 states plus Puerto Rico.

GSRP - Applicants came from 44 states and Puerto Rico. Selectees represent 32 states.

HSAP - Applicants came from thirteen states. Selectees represent nine states.

Total Number of Participants	
SFRP	188
GSRP	109
HSAP	138
TOTAL	435

Degrees Represented			
	SFRP	GSRP	TOTAL
Doctoral	184	1	185
Master's	4	48	52
Bachelor's	0	60	60
TOTAL	188	109	297

SFRP Academic Titles	
Assistant Professor	79
Associate Professor	59
Professor	42
Instructor	3
Chairman	0
Visiting Professor	1
Visiting Assoc. Prof.	0
Research Associate	4
TOTAL	188

Source of Learning About the SRP		
Category	Applicants	Selectees
Applied/participated in prior years	28%	34%
Colleague familiar with SRP	19%	16%
Brochure mailed to institution	23%	17%
Contact with Air Force laboratory	17%	23%
IEEE Spectrum	2%	1%
BIIHE	1%	1%
Other source	10%	8%
TOTAL	100%	100%

APPENDIX B – SRP EVALUATION RESPONSES

1. OVERVIEW

Evaluations were completed and returned to RDL by four groups at the completion of the SRP. The number of respondents in each group is shown below.

Table B-1. Total SRP Evaluations Received

Evaluation Group	Responses
SFRP & GSRPs	275
HSAPs	113
USAF Laboratory Focal Points	84
USAF Laboratory HSAP Mentors	6

All groups indicate unanimous enthusiasm for the SRP experience.

The summarized recommendations for program improvement from both associates and laboratory personnel are listed below:

A. Better preparation on the labs' part prior to associates' arrival (i.e., office space, computer assets, clearly defined scope of work).

B. Faculty Associates suggest higher stipends for SFRP associates.

C. Both HSAP Air Force laboratory mentors and associates would like the summer tour extended from the current 8 weeks to either 10 or 11 weeks; the groups state it takes 4-6 weeks just to get high school students up-to-speed on what's going on at laboratory. (Note: this same argument was used to raise the faculty and graduate student participation time a few years ago.)

2. 1996 USAF LABORATORY FOCAL POINT (LFP) EVALUATION RESPONSES

The summarized results listed below are from the 84 LFP evaluations received.

1. LFP evaluations received and associate preferences:

Table B-2. Air Force LFP Evaluation Responses (By Type)

		How Many Associates Would You Prefer To Get ?								(% Response)			
		SFRP				GSRP (w/Univ Professor)				GSRP (w/o Univ Professor)			
Lab	Evals Recv'd	0	1	2	3+	0	1	2	3+	0	1	2	3+
AEDC	0	-	-	-	-	-	-	-	-	-	-	-	-
WHMC	0	-	-	-	-	-	-	-	-	-	-	-	-
AL	7	28	28	28	14	54	14	28	0	86	0	14	0
FJSRL	1	0	100	0	0	100	0	0	0	0	100	0	0
PL	25	40	40	16	4	88	12	0	0	84	12	4	0
RL	5	60	40	0	0	80	10	0	0	100	0	0	0
WL	46	30	43	20	6	78	17	4	0	93	4	2	0
Total	84	32%	50%	13%	5%	80%	11%	6%	0%	73%	23%	4%	0%

LFP Evaluation Summary. The summarized responses, by laboratory, are listed on the following page. LFPs were asked to rate the following questions on a scale from 1 (below average) to 5 (above average).

2. LFPs involved in SRP associate application evaluation process:
    a. Time available for evaluation of applications:
    b. Adequacy of applications for selection process:
3. Value of orientation trips:
4. Length of research tour:
5 a. Benefits of associate's work to laboratory:
    b. Benefits of associate's work to Air Force:
6. a. Enhancement of research qualifications for LFP and staff:
    b. Enhancement of research qualifications for SFRP associate:
    c. Enhancement of research qualifications for GSRP associate:
7. a. Enhancement of knowledge for LFP and staff:
    b. Enhancement of knowledge for SFRP associate:
    c. Enhancement of knowledge for GSRP associate:
8. Value of Air Force and university links:
9. Potential for future collaboration:
10. a. Your working relationship with SFRP:
    b. Your working relationship with GSRP:
11. Expenditure of your time worthwhile:

(Continued on next page)

12. Quality of program literature for associate:
13. a. Quality of RDL's communications with you:
    b. Quality of RDL's communications with associates:
14. Overall assessment of SRP:

Table B-3. Laboratory Focal Point Reponses to above questions

	AEDC	*AL*	*FJSRL*	*PL*	*RL*	*WHMC*	*WL*
# Evals Recv'd	0	7	1	14	5	0	46
Question #							
2	-	86 %	0 %	88 %	80 %	-	85 %
2a	-	4.3	n/a	3.8	4.0	-	3.6
2b	-	4.0	n/a	3.9	4.5	-	4.1
3	-	4.5	n/a	4.3	4.3	-	3.7
4	-	4.1	4.0	4.1	4.2	-	3.9
5a	-	4.3	5.0	4.3	4.6	-	4.4
5b	-	4.5	n/a	4.2	4.6	-	4.3
6a	-	4.5	5.0	4.0	4.4	-	4.3
6b	-	4.3	n/a	4.1	5.0	-	4.4
6c	-	3.7	5.0	3.5	5.0	-	4.3
7a	-	4.7	5.0	4.0	4.4	-	4.3
7b	-	4.3	n/a	4.2	5.0	-	4.4
7c	-	4.0	5.0	3.9	5.0	-	4.3
8	-	4.6	4.0	4.5	4.6	-	4.3
9	-	4.9	5.0	4.4	4.8	-	4.2
10a	-	5.0	n/a	4.6	4.6	-	4.6
10b	-	4.7	5.0	3.9	5.0	-	4.4
11	-	4.6	5.0	4.4	4.8	-	4.4
12	-	4.0	4.0	4.0	4.2	-	3.8
13a	-	3.2	4.0	3.5	3.8	-	3.4
13b	-	3.4	4.0	3.6	4.5	-	3.6
14	-	4.4	5.0	4.4	4.8	-	4.4

3. 1996 SFRP & GSRP EVALUATION RESPONSES

The summarized results listed below are from the 257 SFRP/GSRP evaluations received.

Associates were asked to rate the following questions on a scale from 1 (below average) to 5 (above average) - by Air Force base results and over-all results of the 1996 evaluations are listed after the questions.

1. The match between the laboratories research and your field:
2. Your working relationship with your LFP:
3. Enhancement of your academic qualifications:
4. Enhancement of your research qualifications:
5. Lab readiness for you: LFP, task, plan:
6. Lab readiness for you: equipment, supplies, facilities:
7. Lab resources:
8. Lab research and administrative support:
9. Adequacy of brochure and associate handbook:
10. RDL communications with you:
11. Overall payment procedures:
12. Overall assessment of the SRP:
13. a. Would you apply again?
    b. Will you continue this or related research?
14. Was length of your tour satisfactory?
15. Percentage of associates who experienced difficulties in finding housing:
16. Where did you stay during your SRP tour?
    a. At Home:
    b. With Friend:
    c. On Local Economy:
    d. Base Quarters:
17. Value of orientation visit:
    a. Essential:
    b. Convenient:
    c. Not Worth Cost:
    d. Not Used:

SFRP and GSRP associate's responses are listed in tabular format on the following page.

Table B-4. 1996 SFRP & GSRP Associate Responses to SRP Evaluation

	Arnold	Brooks	Edwards	Eglin	Griffis	Hanscom	Kelly	Kirtland	Lackland	Robins	Tyndall	WPAFB	average
# res	6	48	6	14	31	19	3	32	1	2	10	85	257
1	4.8	4.4	4.6	4.7	4.4	4.9	4.6	4.6	5.0	5.0	4.0	4.7	4.6
2	5.0	4.6	4.1	4.9	4.7	4.7	5.0	4.7	5.0	5.0	4.6	4.8	4.7
3	4.5	4.4	4.0	4.6	4.3	4.2	4.3	4.4	5.0	5.0	4.5	4.3	4.4
4	4.3	4.5	3.8	4.6	4.4	4.4	4.3	4.6	5.0	4.0	4.4	4.5	4.5
5	4.5	4.3	3.3	4.8	4.4	4.5	4.3	4.2	5.0	5.0	3.9	4.4	4.4
6	4.3	4.3	3.7	4.7	4.4	4.5	4.0	3.8	5.0	5.0	3.8	4.2	4.2
7	4.5	4.4	4.2	4.8	4.5	4.3	4.3	4.1	5.0	5.0	4.3	4.3	4.4
8	4.5	4.6	3.0	4.9	4.4	4.3	4.3	4.5	5.0	5.0	4.7	4.5	4.5
9	4.7	4.5	4.7	4.5	4.3	4.5	4.7	4.3	5.0	5.0	4.1	4.5	4.5
10	4.2	4.4	4.7	4.4	4.1	4.1	4.0	4.2	5.0	4.5	3.6	4.4	4.3
11	3.8	4.1	4.5	4.0	3.9	4.1	4.0	4.0	3.0	4.0	3.7	4.0	4.0
12	5.7	4.7	4.3	4.9	4.5	4.9	4.7	4.6	5.0	4.5	4.6	4.5	4.6
Numbers below are percentages													
13a	83	90	83	93	87	75	100	81	100	100	100	86	87
13b	100	89	83	100	94	98	100	94	100	100	100	94	93
14	83	96	100	90	87	80	100	92	100	100	70	84	88
15	17	6	0	33	20	76	33	25	0	100	20	8	39
16a	-	26	17	9	38	23	33	4	-	-	-	30	
16b	100	33	-	40	-	8	-	-	-	-	36	2	
16c	-	41	83	40	62	69	67	96	100	100	64	68	
16d	-	-	-	-	-	-	-	-	-	-	-	0	
17a	-	33	100	17	50	14	67	39	-	50	40	31	35
17b	-	21	-	17	10	14	-	24	-	50	20	16	16
17c	-	-	-	-	10	7	-	-	-	-	-	2	3
17d	100	46	-	66	30	69	33	37	100	-	40	51	46

4. 1996 USAF LABORATORY HSAP MENTOR EVALUATION RESPONSES

Not enough evaluations received (5 total) from Mentors to do useful summary.

5. 1996 HSAP EVALUATION RESPONSES

The summarized results listed below are from the 113 HSAP evaluations received.

HSAP apprentices were asked to rate the following questions on a scale from 1 (below average) to 5 (above average)

1. Your influence on selection of topic/type of work.
2. Working relationship with mentor, other lab scientists.
3. Enhancement of your academic qualifications.
4. Technically challenging work.
5. Lab readiness for you: mentor, task, work plan, equipment.
6. Influence on your career.
7. Increased interest in math/science.
8. Lab research & administrative support.
9. Adequacy of RDL's Apprentice Handbook and administrative materials.
10. Responsiveness of RDL communications.
11. Overall payment procedures.
12. Overall assessment of SRP value to you.
13. Would you apply again next year? Yes (92 %)
14. Will you pursue future studies related to this research? Yes (68 %)
15. Was Tour length satisfactory? Yes (82 %)

	Arnold	Brooks	Edwards	Eglin	Griffiss	Hanscom	Kirtland	Tyndall	WPAFB	Totals
# resp	5	19	7	15	13	2	7	5	40	113
1	2.8	3.3	3.4	3.5	3.4	4.0	3.2	3.6	3.6	3.4
2	4.4	4.6	4.5	4.8	4.6	4.0	4.4	4.0	4.6	4.6
3	4.0	4.2	4.1	4.3	4.5	5.0	4.3	4.6	4.4	4.4
4	3.6	3.9	4.0	4.5	4.2	5.0	4.6	3.8	4.3	4.2
5	4.4	4.1	3.7	4.5	4.1	3.0	3.9	3.6	3.9	4.0
6	3.2	3.6	3.6	4.1	3.8	5.0	3.3	3.8	3.6	3.7
7	2.8	4.1	4.0	3.9	3.9	5.0	3.6	4.0	4.0	3.9
8	3.8	4.1	4.0	4.3	4.0	4.0	4.3	3.8	4.3	4.2
9	4.4	3.6	4.1	4.1	3.5	4.0	3.9	4.0	3.7	3.8
10	4.0	3.8	4.1	3.7	4.1	4.0	3.9	2.4	3.8	3.8
11	4.2	4.2	3.7	3.9	3.8	3.0	3.7	2.6	3.7	3.8
12	4.0	4.5	4.9	4.6	4.6	5.0	4.6	4.2	4.3	4.5
Numbers below are percentages										
13	60%	95%	100%	100%	85%	100%	100%	100%	90%	92%
14	20%	80%	71%	80%	54%	100%	71%	80%	65%	68%
15	100%	70%	71%	100%	100%	50%	86%	60%	80%	82%

INDIVIDUAL DIFFERENCES IN DUAL -TASK PERFORMANCE: EFFECTS OF HANDEDNESS AND FAMILIAL SINISTRALITY ON THE ABILITY TO FLY A SIMULATED AIRPLANE

Ruth E. Propper
Graduate Student
Department of Psychology

University of Toledo
Toledo, Ohio
43606

Final Report for:
Summer Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

September, 1996

INDIVIDUAL DIFFERENCES IN DUAL -TASK PERFORMANCE: EFFECTS OF HANDEDNESS AND FAMILIAL SINISTRALITY ON THE ABILITY TO FLY A SIMULATED AIRPLANE

Ruth E. Propper
Graduate Student
Department of Psychology
University of Toledo

Abstract

The influence of familial sinistrality (FS) and degree of subjects' handedness on the ability to perform in a dual-task situation that simulated the control of pitch, roll, and yaw in an airplane's cockpit was examined. While FS and subjects' personal handedness did exert an effect on the dependent measures, the fact that results did not attain significance makes interpretation difficult. Factors limiting the generalizability of the results are discussed, and suggestions for future research proposed. In addition, a tentative explanation for significant ($p < .01$) hand use effects is offered, with the suggestion that greater left hand dual-task interference during the performance of a unimanual spatial task takes place at the level of spatial processing, while greater right hand dual-task interference during performance of a bipedal spatial task takes place at the level of response.

INDIVIDUAL DIFFERENCES IN DUAL -TASK PERFORMANCE: EFFECTS OF HANDEDNESS AND FAMILIAL SINISTRALITY ON THE ABILITY TO FLY A SIMULATED AIRPLANE

Ruth E. Propper

Introduction

In order to fly an airplane, a pilot must be able to perform multiple tasks simultaneously and with a minimum of between-task interference. For example, a pilot must be able to monitor and control the movement of the plane while concurrently listening for communications from mission control and from other aircraft, attending to fuel levels, and monitoring engine performance (Ikomi & Tirre, 1994). In addition, aircraft control itself involves the performance of multiple tasks; pilots must control airplane pitch (movement about the lateral axis), roll (movement about the longitudinal axis), and yaw (movement about the vertical axis). In most aircraft, indices and control of pitch and roll are combined in a single visual display and rely on a single response mechanism, while control of yaw is distinct from these other forms of motion both visually and in the form of response required. Because two different displays and forms of response are necessary to maintain control over the motion of an aircraft, keeping a plane on a particular flight attitude is a dual-task situation.

It is has been suggested that dual-task decrements in performance (a decrease in ability from single- to multiple-task situations) are the result of competition between tasks for particular resources (Friedman, Polson, & Dafoe, 1988; Kinsbourne & Hicks, 1978; Navon & Gopher, 1979; Wickens, 1980, 1984, 1991, 1992), with resources being defined variously in terms of processing stages, codes of processing, input modality (See Wickens, 1991, 1992), as well as in structure (Friedman & Polson, 1981; Friedman, Polson & Dafoe, 1988; Kinsbourne & Hicks, 1978; Kinsbourne & Hiscock, 1983). Generally, to the extent that two concurrently performed tasks compete for the same resources, decreases in performance will be predicted to occur (See Fracker & Wickens, 1989; Lane, 1982 for more detailed discussions of resource theory).

For example, Wickens, Mountford, and Schreiner (1981) reported that a visual-spatial tracking task performed in conjunction with an auditory memory task resulted in less single- to dual-task decline than did a tracking task performed simultaneously with either of two other spatial tasks. In the tracking task, subjects, used a joystick to keep a cursor centered on a moving verticle bar presented on a computer screen, thereby heavily engaging spatial, visual, and manual response resources. In contrast, the auditory memory task utilized primarily verbal and auditory resources; subjects were required to judge whether letters presented through headphones were alphabetically sequential or nonsequential, and were told to respond by pressing a button on a keyboard to indicate their choice. Because dual-task performance did not suffer as much when these two tasks were performed simultaneously compared to when the tracking task was paired with either of two other spatial tasks, it was suggested that dividing processing requirements between resources increased the ability to perform two tasks simultaneously, or to time-share.

It has been suggested that the cerebral hemispheres represent two structurally defined resources (Friedman & Polson, 1981; Friedman, Polson, & Dafoe, 1988; Kinsbourne & Cook, 1971; Wickens, 1980; Wickens, Mountford, & Schreiner, 1981). According to these hypotheses, dual-task decrements in performance may sometimes reflect intrahemispheric competition for information processing resources (Friedman & Polson, 1981; Friedman, Polson, & Dafoe, 1988; Kinsbourne & Cook, 1971). For example, the ability to balance a dowel rod with the right hand decreases when subjects are required to speak while performing the task. Left-hand dowel balancing does not show this dual-task decrement (Kinsbourne & Cook, 1971). Because the left cerebral hemisphere is responsible for both language and motor movement of the right hand, during concurrent right-hand dowel balancing and speech production, the two compete for the same resources within a hemisphere, resulting in task interference. During concurrent performance of left-hand dowel balancing and speech, however, processing requirements are divided between the hemispheres. The fact that dual-task decrements in performance increase with increasing demands on a given hemisphere, while dividing the processing requirements of two tasks between the hemispheres reduces dual-task interference, has been used to support the hypothesis that intrahemispheric

competition for resources may result in dual-task interference (Friedman & Polson, 1981; Friedman, Polson & Dafoe, 1988; Kinsbourne & Cook, 1971; Kinsbourne & Hicks, 1978; Kinsbourne & Hiscock, 1983).

The left hemisphere is primarily responsible for processing verbal stimuli, while the right hemisphere is primarily responsible for processing spatial material. (Bradshaw & Nettleton, 1981). According to a model that proposes that the two cerebral hemispheres represent structurally defined resources, a spatial task performed in conjunction with a verbal task should result in less dual-task interference than two spatial tasks performed concurrently, and research has supported this prediction (Wickens, Mountford, & Schreiner, 1981). In the experiment conducted by Wickens et al., it was found that performance on a spatial tracking task suffered more when it was paired with either of two other spatial tasks, compared to when it was performed simultaneously with a verbal task. Although the results were discussed in terms of Wicken's (1980) multiple resource model of dual-task interference, it is possible that the obtained results reflect degrees of intrahemispheric competition for resources.

In the present experiment, control of both pitch/roll and yaw involve tracking visually displayed stimuli and spatially processing the information presented in order to respond. Because both tasks are competing for right hemisphere spatial information processing resources, a dual-task decrement in performance is predicted to occur. If intrahemispheric competition for right hemisphere resources can result in poor dual-task performance, then the extent to which spatial processing itself is lateralized to the right cerebral hemisphere may influence the ability to perform two spatial tasks simultaneously. For example, distribution of spatial information processing resources, or of the neural substrates underlying spatial ability, between the hemispheres, could result in less dual-task interference during concurrent performance of two spatial tasks; task processing requirements could be divided between the hemispheres, reducing intrahemispheric competition and lessening the dual-task decrement.

There are a group of individuals who have been demonstrated to have less lateralization of hemispheric functions than do other individuals. Left- and weakly-handed individuals (those with no strong hand use preference), and people with familial sinistrality (FS+), appear to have less hemispheric lateralization of function than do strongly- and/or right-handed individuals, or those without familial sinistrality (FS-) (see discussions by

Beaton, 1985 and McKeever, 1990). With particular relevance to the present experiment, research has been reported demonstrating that left/weakly-handed individuals, or those who are FS+, tend to have more bilateral representation of spatial functions than do individuals who are right/strongly-handed, or who are FS- (Hardyk, 1977; Hecaen, DeAgostini, & Monzon-Montes, 1981; Marino & McKeever, 1989; McKeever, 1990; Schmuller & Goodman, 1980; Piazza 1980)

For example, Hecaen, DeAgostini, & Monzon-Montes (1981) examined the incidence of spatial disorders in 271 left- and right-handed subjects with left and right hemisphere lesions. In right handers, right hemisphere lesions resulted in a significantly higher incidence of spatial disorder than left hemisphere lesions. In contrast, left and right hemisphere lesions resulted in no difference in the incidence of spatial disorders in left-handers. It can be suggested, based on these results, that the left and right cerebral hemispheres are equally capable of processing spatial stimuli in left-handed individuals.

It is possible that left/weak-handers, or those who are FS+ will have less of a dual-task performance decrement, compared to strong/right handers, when performing pitch/roll and yaw simultaneously. Since both tasks require spatial processing, it is predicted that right/strongly-handed individuals and those who are FS- will experience greater dual task performance deficits as a result of intrahemispheric competition between tasks. Because spatial processing resources are more bilaterally distributed in left/weak-handers, and those who are FS+, these individuals may be better able to divide task processing requirements between the cerebral hemispheres, resulting in a lesser decline in performance during the dual-task condition.

Although all subjects are predicted to perform more poorly on both pitch/roll and yaw when these tasks are performed simultaneously, compared to when they are performed alone, it is predicted that left/weak-handers and those who are FS+ will demonstrate less of a dual-task decrement than will right/strong-handers and those who are FS-. In addition, decrements in performance of pitch/roll under dual-task conditions may be asymmetrical as a function of handedness (and/or FS) and hand used for responding (management of pitch/roll requires unimanual responding). In right-handers, motor control of the limbs is controlled by the hemisphere contralateral to the limb. Left-hand responding in right-handed, FS- individuals will therefore compete for right hemisphere

resources, resulting in a decrease in performance, compared to when these subjects respond with their right hand. Because both spatial abilities and motor control of the hands is more bilaterally distributed in left-handers (Honda, 1982), dual-task decrements in the performance of pitch/roll should be equivalent between the hands in these subjects. In fact, there is evidence that dual-task decrements in performance in left/weak-handers, and those who are FS+ do tend to be less asymmetrical between hand of response than in right/strong-handers, or those who are FS-, although this effect is not always found (see Hellige & Kee, 1990; Hiscock, Kinsbourne, & Green, 1990 for discussions on the use of the dual-task paradigm and its implications for cerebral lateralization of function in left/weak-handers).

Control of yaw involves responding with both feet by depressing foot pedals. Because, to the authors' knowledge, there has been no research examining dual-task performance when responding involves the simultaneous use of both lower limbs in conjunction with unimanual responding, the only prediction that can be made with regard to yaw is that dual-task performance will be inferior to single-task ability.

It has been suggested that the ability to perform two tasks simultaneously increases with practice on either the single- or dual-task trials. The improvement in subjects' ability to perform two tasks concurrently can, in part, reflect the increased automaticity, and therefore the decreased resource demands, of the single tasks (Wickens, 1992). If decreased resource demands are responsible for increases in dual-task performance, it is possible that individual differences in performance will diminish with training. Because, by lessening the resource demands of a task, the effects of intrahemispheric competition for resources will no longer be an issue in dual-task ability, individual differences in cerebral lateralization on dual-task performance will be minimized with practice. In addition, it is possible that with experience subjects will develop strategies to help them perform yaw and pitch/roll simultaneously. These strategies may also minimize the influence of individual differences on performance.

Method

Participants

Participants were 153 male USAF recruits, on their 32nd or 33rd day of basic training at Lackland Air Force Base, San Antonio, Texas. All subjects were high school graduates. Participants' handedness and familial sinistrality (FS) was assessed via the Edinburgh Handedness Inventory (Oldfield, 1971), with subjects scoring +70 and above, or -70 and below, considered strongly handed, and those scoring between these two extremes considered weakly handed.

Apparatus and Stimuli

All tests and instructions were presented on either Pentium computers (12 90 MHZ and 10 166 MHZ) or 486 DX computers (3 Goldstar, 10 Antares, and 10 Dell) equipped with 17 inch color monitors, CH product flight sticks, and Pro Pedal CH Product rudder pedals.

Yaw Stimuli: A yellow ball, approximately 8 pixels by 12 pixels, inside of a rectangle approximately 22 pixels by 131 pixels, moved left and right randomly inside the rectangle. Ten hash marks, five to either side of the center of the rectangle, spaced equally apart, could be used as cues by the subject to determine how far from center the ball was located. Subjects' goal was to keep the ball centered by depressing the rudder pedals. Pressing the right rudder pedal caused the ball to move to the left, while pressing the left pedal caused the ball to move to the right. During both single- and dual-task trials the display was centered horizontally in the middle of the computer screen, and approximately 2 degrees vertically below the center of the screen.

Roll and Pitch Stimuli: The display for both roll and pitch consisted of a simulated horizon, made by coloring half of a circle, approximately 69 pixels by 96 pixels, brown to suggest land, and half blue, to suggest sky. The horizon itself was a darkened line. Miniature airplane wings spanned the diameter of the circle, and included a small airplane tail over the center of the display. In roll, the simulated horizon could rotate, clockwise or counter-clockwise, 360 degrees. Subjects' goal was to keep the horizon horizontal within the display by manipulating the joystick; left lateral movements caused the horizon to rotate counter-clockwise, while right lateral movements caused the horizon to rotate clockwise, giving the subject the illusion that he/she was controlling the movement of the airplane.

The display used for pitch was identical to that used for roll. In pitch, however, the horizon could move vertically, either toward the top or bottom of the display. Subjects' goal was to keep the airplane wings centered on the horizon; pushing the joystick forward resulted in the horizon moving to the bottom of the display, while pulling the joystick back resulted in the horizon moving toward the top of the display (see Figures 1 to 4b for depictions of stimuli).

Random displacement of the horizon for pitch/roll, and of the ball for yaw, occurred at a rate of 1 pixel per 14 msec. Direction of movement was randomly determined every five seconds, with stimuli either continuing in the same direction as in the previous five seconds, changing directions, or remaining motionless. Subjects' compensation for random stimuli movement was determined by how hard the rudder pedals were pushed for yaw, or the flight stick was moved for pitch/roll, and ranged from 3 to 0 pixels per 14 msec.

During two brief (15 sec) orientation trials, roll and pitch were performed separately. For the remainder of the experiment however, pitch and roll were combined and performed simultaneously, using one joystick and one display. During dual-task trials, the yaw and pitch/roll displays were centered on the computer screen in a verticle cluster, with the yaw indicator below the pitch/roll display.

<u>Procedure</u>

As part of a larger test battery, subjects were tested in groups in 6 experimental sessions, seated at individual work stations equipped with the necessary apparatus. Other tests included in the battery were several cognitive tasks and four questionnaires. Results of these other tests will be reported elsewhere and are not relevant to the present experiment.

Before beginning the task proper, subjects completed the orientation trials. No hand use instructions were given for these trials, and it is probable that subjects simply used their dominant hand. Following orientation trials, subjects completed five 5 minute test trials; three single-task trials, and two dual-task trials. Single-task trials consisted of; yaw alone, pitch/roll alone with the subjects' right hand, and pitch/roll alone with the left hand. Order of single-task trials was counterbalanced between subjects. After completing all single-task trials, subjects performed dual-task trials. These consisted of pitch/roll and yaw performed simultaneously, with

pitch/roll controlled with either the right or left hand. Dual-task order was counterbalanced between subjects. Subjects could take a brief break between all trial types. Instructions regarding which hand the subject was to use were displayed on the computer screen at the beginning of any block that contained pitch/roll trials (excluding orientation trials). Although only one joystick per subject was available, the joystick could be transferred to either hand depending on task requirements. Experimenters were available to aid subjects if confusion arose.

Following the final set of trials in the first block, subjects performed two other tests in the battery. These tasks lasted approximately 45 minutes. After completing these tasks, subjects were again presented with the 5 test trials, although this second time they were presented in a random order. The entire testing session lasted approximately 3 hours.

Results and Discussion

A subjects' performance was based on the average distance he/she was from the target location (ie: the center of the display for yaw, on the horizon for pitch, and flying level with the horizon for roll). For scoring purposes, the displays were divided into 17 segments, ranging from -1 to -9 and from +1 to +9. Zero was considered the center, and indicated that the subject was on target (see Figure 1). In roll, if the subject was flying level with the horizon, he/she received a score of zero. Counter-clockwise roll of the horizon to the miniature airplane wings (indicating the subject flew primarily upside-down) was scored -1 to -9. Clockwise roll of the horizon to the miniature airplane wings was scored +1 to +9 (see Figures 2a and 2b). In pitch, if the subject was flying straight toward the ground (ie, the display was completely brown), he/she received a score of -9. On the other hand, if the subject was flying straight toward the sky, he/she received a score of +9. If the plane was in any way touching the horizon, the subject received a score of zero (see Figures 3a and 3b). It was possible for subjects to receive a zero score for pitch, while simultaneously receiving a score of -9 for roll, and vice-versa. For yaw, the display was again divided into 17 segments, with the center position being considered zero, left of center scored -1 to -9, and right of center scored +1 to +9 (see Figures 1, 4a, and 4b).

For analysis, subjects' position readings were calculated every 14 msec and were averaged over entire five minute blocks. For ease of interpretation, subjects' absolute scores were subtracted from nine; higher scores

therefore indicated better performance. Analysis was conducted on dual-task decrement scores that followed the formula (9- | single-task performance |) - (9 - | dual-task performance |). A 2 (FS; + or -) X 2 (Handedness; strong or weak) X 2 (Hand; left or right) X 2 (Block; first or second) MANOVA was conducted on these pitch, roll, and yaw dual-task decrement scores.

Results revealed a trend for a main effect of FS on performance of pitch, $\underline{F}$ (1, 153) = 2.77, p = .098, with FS+ subjects (n = 74) showing a greater dual-task decrement (M = .49) compared to FS- subjects (n = 79, M = .28). Examination of absolute raw scores revealed that this effect was due primarily to the tendency of FS+ subjects to exhibit superior single-task performance (M = 8.87), compared to FS- subjects (M = 8.78), while both groups demonstrated equivalent dual-task performance (M = 8.66 and M = 8.65, respectively). That is, although FS+ subjects demonstrated a greater dual-task decrement, this effect was not due to these subjects suffering from greater dual-task interference, compared to FS- subjects, but rather to FS+ superior single-task ability. This effect is contrary to reports demonstrating decreased spatial processing ability in FS+ individuals (see McKeever, 1990, for a review). In light of the fact that the results did not attain significance in either absolute raw or decrement scores, these results cannot be interpreted with any degree of certainty, and require replication.

Analysis of roll decrement scores revealed a main effect of Hand, $\underline{F}$ (1, 153) = 12.19, p < .01, with subjects demonstrating a greater dual-task decrement using their left hand (M = .16) compared to their right hand (M = .11). There was a trend for a four-way interaction of Handedness, FS, Hand, and Practice, $\underline{F}$ (1, 153) = 3.2, p = .076, with FS+, weakly-handed individuals (n = 31) demonstrating the least dual-task decrement during Block 1 presentations (M = .08), and FS-, strongly-handed individuals (n = 44) demonstrating the greatest (M = .21). In addition, FS-, strongly-handed subjects demonstrated the greatest left-right hand difference in decrement scores (M = .08) during Block 1 presentations. While these results are in the predicted direction, decrement scores were extremely small and again did not attain significance, and therefore any interpretations or generalizations must be accepted with caution (see Figures 5a and 5b).

Results revealed a main effect of Hand on yaw decrement scores, $\underline{F}$ (1, 153) = 11.81, $p < .001$, with subjects experiencing a greater dual-task decrement when they used their right hand to perform pitch/roll (M = .84), compared to their left (M = .68). There was a trend for a main effect of FS, $\underline{F}$ (1, 153) = 3.04, p = .084, with FS- subjects demonstrating a greater dual-task decrement (M = .86), compared to FS+ subjects (M = .67). Examination of absolute raw scores revealed that this effect was indeed due to FS- subjects' poorer performance on dual-task trials (M = 7.83), compared to FS+ subjects (M = 7.99). FS+ and FS- subjects performed comparably on single-task trials (M = 8.64, M = 8.69, respectively). Again, these results must be interpreted with caution, and require replication.

It should be noted that subjects performed extremely well on both single- and dual-task trials, with single-task scores nearly perfect (pitch M = 8.78, roll M = 8.78, yaw M = 8.67), and dual-task scores quite close as well (pitch M = 8.66, roll M = 8.65, yaw M = 7.91). (Although analysis did reveal performance on dual-task trials to be significantly worse than single-task trials across all dependent measures, $p < .01$). Hand and FS effects are typically quite small and variable, and it is likely that any between-subject effects or interactions were masked by ceiling effects.

Furthermore, although performance of yaw suffered more under dual-task conditions than did pitch or roll, it is probable that this effect is an artifact of the order of initial single-task presentations; it is possible that those subjects who received yaw single-task trials first perceived this task as the primary one, sacrificing pitch/roll, while those subjects who initially received pitch/roll sacrificed yaw performance. The fact that pitch/roll trials were two times more likely to be shown first than was yaw could account for the greater dual-task decrement found for this dependent variable. Variability in subjects' performance was therefore increased due to the lack of any primary/secondary task manipulations, further masking any potential effects of handedness or FS.

With these caveats in mind, an explanation of the relatively robust Hand main effects found for yaw and roll may be tentatively offered. Ipsilateral, as well as contralateral, cortical pathways exist for the control of gross motor movements generally, and for lower limb movements in particular. It is therefore possible that subjects' left hemisphere took responsibility for the bipedal responses required of yaw in order to lessen right hemisphere

processing load. Right-hand control of pitch/roll, dependent on left-hemisphere processes, would therefore interfere with yaw at the level of the response, resulting in a greater dual-task decrement in yaw, compared to when subjects used their left hands. This effect would be particularly pronounced if, as mentioned above, 2/3 of subjects believed pitch/roll to be the primary task. Left hand interference of right-hemisphere spatial processes would result in the greater dual-task decrement found when subjects used their left hand for roll. Although these effects would be predicted to interact with subjects' handedness and familial sinistrality, as already mentioned, between subject effects would have been masked by ceiling effects and by variability in subjects' perception of the primary task.

Future research could address these issues by increasing task difficulty or by manipulating the primary task. For example, if dual-task interference of bipedal responses is due to intrahemispheric competition for motor response resources, then suggestions that performance on a bipedal task is most important should result in a greater or equivalent right hand, compared to left hand, dual-task decrement for the performance of a unimanual task, at least in strongly-handed, or FS- subjects. Increasing stimuli displacement would increase task difficulty. This could be accomplished by decreasing the length of time between stimuli direction changes, or by eliminating those moments when stimuli remain motionless.

References

Beaton, A. (1985). Left side, right side: A review of laterality research. New Haven: Yale University Press.

Bradshaw, J. L., & Nettleton, N. C. (1981). The nature of hemispheric specialization in man. The Behavioral and Brain Sciences, 4, 51-91.

Fracker, M. L., & Wickens, C. D. (1989). Resources, confusions, and compatibility in dual-axis tracking: Displays, controls, and dynamics. Journal of Experimental Psychology: Human Perception and Performance, 15, 80-96.

Friedman, A. & Polson, M. C. (1981). The hemispheres as independent resource systems: Limited capacity processing and cerebral specialization. Journal of Experimental Psychology: Human Perception and Performance, 7, 1031-1058.

Friedman, A., Polson, M.C., & Dafoe, C. G. (1988). Dividing attention between the hands and the head: Performance trade-offs between rapid finger tapping and verbal memory. Journal of Experimental Psychology: Human Perception and Performance, 14, 60-68.

Hardyck, C. (1977). A model of individual differences in hemispheric specialization. In H. Whitaker & H. A. Whitaker (Eds.), Studies in Neurolinguistics (vol. 3). New York: Academic Press.

Hecaen, H., DeAgostini, M., & Monzon-Montes, A. (1981). Cerebral organization in left-handers. Brain and Language, 12, 261-284.

Hellige, J. B., & Kee, D. W. (1990). Asymmetric manual interference as an indicator of lateralized brain function. In G. E. Hammond (Ed.), Cerebral control of speech and limb movements (pp. 636-660). New York: Elsevier Science Publishers B. V. (North- Holland).

Hiscock, M., Kinsbourne, M., & Green, A. (1990). Is time-sharing asymmetry a valid indicator of speech lateralization? Evidence for left-handers. In G. E. Hammond (Ed.), Cerebral control of speech and limb movements (pp. 611-633). New York: Elsevier Science Publishers B.V. (North-Holland).

Honda, H. (1982). Effect of handedness on the inhibition of reaction time in double stimulation situations. Japanese Psychological Review, 24, 43-47.

Ikomi, P. A., & Tirre, W. (1994). A realistic multi-task assessment of pilot aptitudes. (Summer Faculty Research Program Final Report). Bolling, AFB, DC: Air Force Office of Scientific Research.

Kinsbourne, M., & Cook, J. (1971). Generalized and lateralized effects of concurrent verbalization on a unimanual skill. Quarterly Journal of Experimental Psychology, 23, 341-345.

Kinsbourne, M., & Hicks, R. E. (1978). Functional cerebral space: A model for overflow, transfer, and interference effects in human performance. In J. Requin (Ed.), Attention and Performance VII (pp. 345-362). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kinsbourne, M., & Hiscock, M. (1983). Asymmetrics in dual task performance. In J. Hellige (Ed.), Cerebral hemisphere asymmetry: Method, theory, and application (pp. 255-344). New York: Praeger.

Lane, D. M. (1982). Limited capacity, attention allocation, and productivity. In W. C. Howell & E. A. Fleishman (Eds.), Human performance and productivity: Information processing and decision making (Vol. 2, pp. 121-156). Hillsdale, NJ: Lawrence Erlbaum Associates.

Marino, M. F., & McKeever, W. F. (1989). Spatial processing laterality and spatial visualization ability: Relations to sex and familial sinistrality. Bulletin of the Psychonomics Society, 27, 135-137.

McKeever, W. F. (1990). Familial sinistrality and cerebral organization. In S. Coren (Ed.), Behavior implications and anomalies (pp. 373-412). New York: Elsevier Science Publishers B. V. (North-Holland).

Navon, D., & Gopher, D. (1979). On the economy of the human processing system. Psychological Review, 86, 214-225.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. Neuropsychologia, 9, 97-111.

Piazza, D. M. (1980). The influence of sex and handedness in the hemispheric specialization of verbal and nonverbal tasks. Neuropsychologia, 18, 163-176.

Schmuller, J., & Goodman, R. (1980). Bilateral tachistoscopic perception, handedness, and laterality. Brain and Language, 11, 12-18.

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), Attention and performance VIII. Hillside, NJ: Erlbaum.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), Varieties of attention (pp. 63-102). New York: Academic Press.

Wickens, C. (1991). Processing resources and attention. In D. L. Damos (Ed.), Multiple-task performance (pp. 3-34). Washington, DC: Taylor and Francis.

Wickens, C. D. (1992). Attention, time-sharing, and workload. In C. D. Wickens (Ed.), Engineering psychology and human performance (2nd ed., pp. 364-411). Champaign-Urbana, Illinois: Harper-Collins.

Wickens, C. D., Mountford, S. J., & Schreiner, W. (1981). Multiple resources, task-hemispheric integrity, and individual differences in time-sharing. Human Factors, 23, 211-229.

Figure Captions

Figure 1. Target positions of pitch/roll and yaw stimuli displays.

Figures 2a and 2b. Examples of horizon displacement and scoring during roll, with numbers in parentheses indicating subject's score.

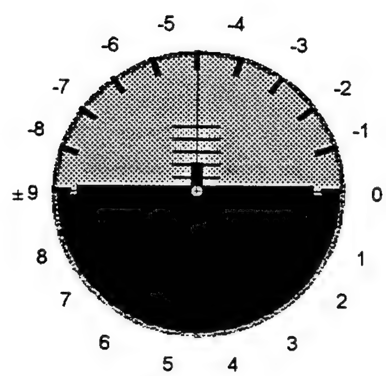
Figures 3a and 3b. Examples of horizon displacement and scoring during pitch, with numbers in parentheses indicating subject's score.

Figures 4a and 4b. Examples of horizon, ball displacement and scoring during dual-task trials. Numbers in parentheses indicate subject's score.

Figures 5a and 5b. Roll dual-task decriment scores as a function of FS, Handedness, Hand, and Practice.

Figure 1

Pitch & Roll Indicator



Yaw Indicator

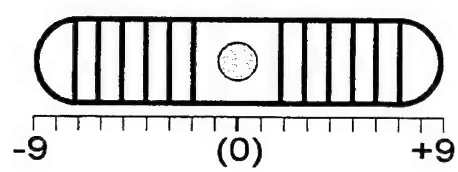


Figure 2

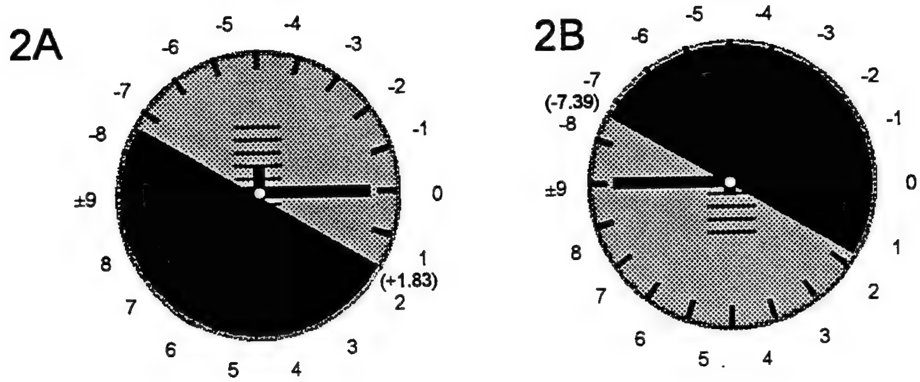


Figure 3

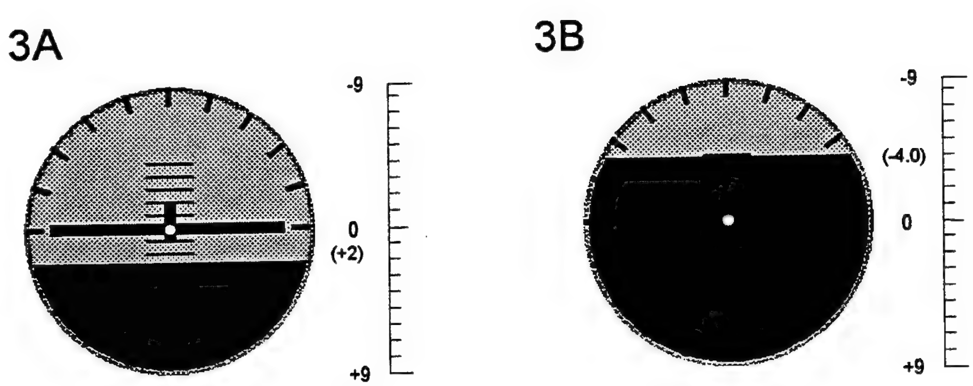
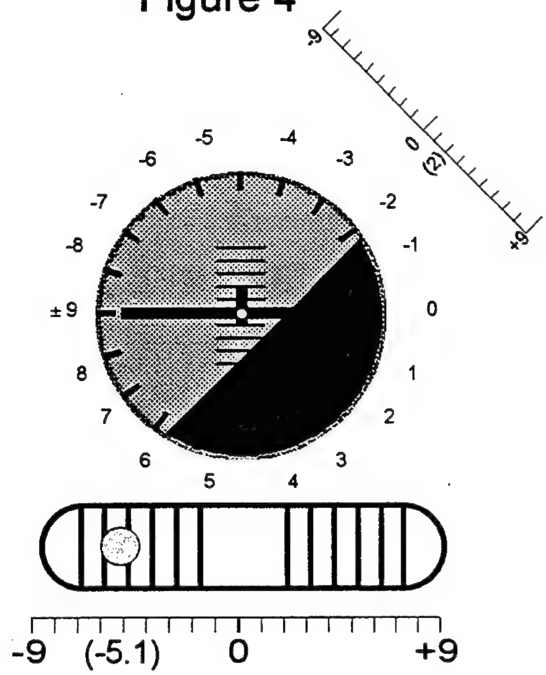


Figure 4

4A



4B

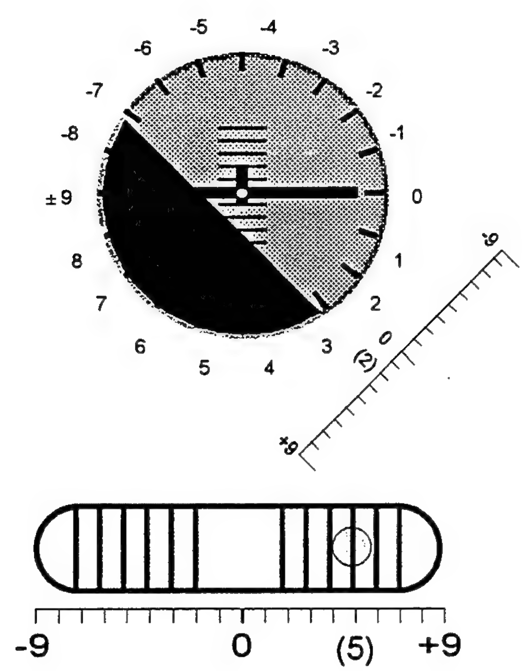
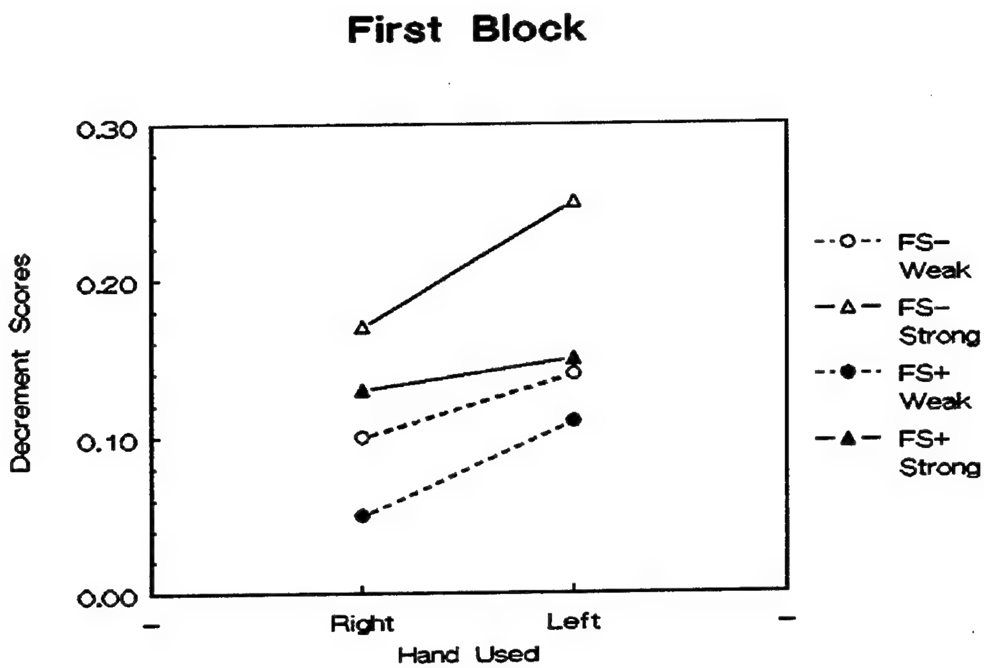
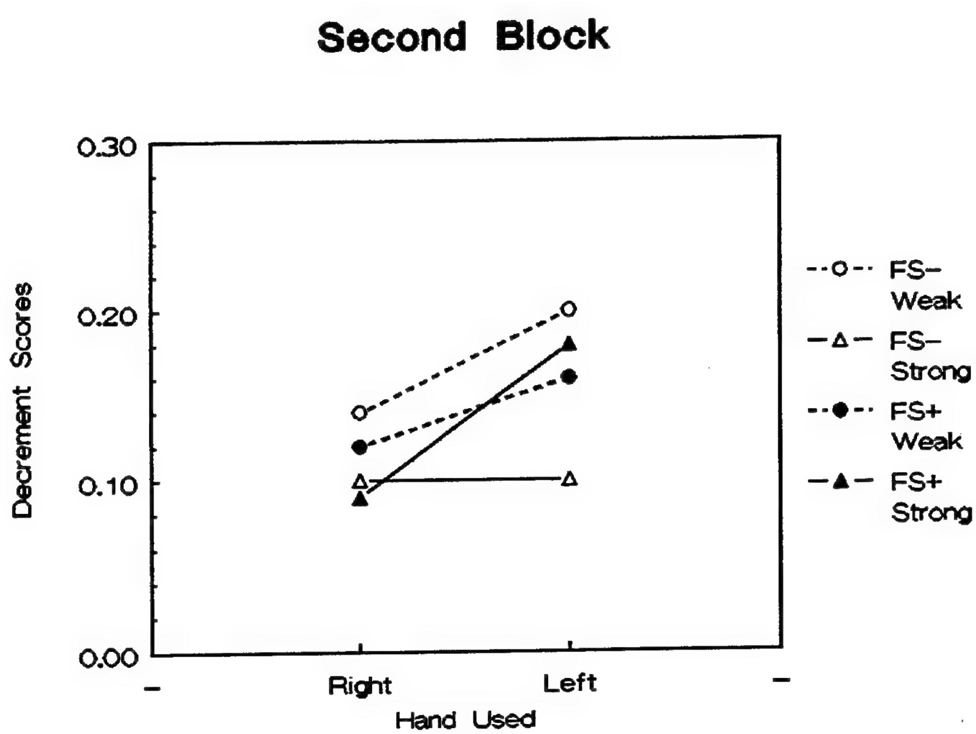


Figure 5

5A



5B



Detection of Escherichia Coli by Multiplex Polymerase Chain Reaction

Catherine A. Ramaika
Graduate Student
University of Texas School of Public Health
San Antonio, Texas 78284

Final Report for:
Summer Graduate Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Brooks Air Force Base, Texas

and

Armstrong Laboratory

August 1996

Detection of Escherichia coli O157:H7 by Multiplex Polymerase Chain Reaction

Catherine A. Ramaika
Graduate Student
University of Texas School of Public Health
San Antonio, Texas 78284

Abstract

Escherichia coli O157:H7 is a new enterohemorrhagic (EHEC) pathogen that has become an important public health issue since it has been implicated in recent food-borne outbreaks due to incompletely cooked ground beef. The ability to identify O157:H7 from other E. coli species and organisms with multiplex PCR conditions would enable a quick and specific diagnosis for this organism. This study tested a number of samples for a 60 MDa plasmid and for Shiga-like toxins (I & II) for PCR conditions. Once established that the test was accurate, multiplex conditions were determined for both plasmid and toxins.

Detection of <u>Escherichia Coli</u> by Multiplex Polymerase Chain Reaction

<u>INTRODUCTION</u>

The United States experienced a foodborne disease outbreak in 1982 which resulted in the recognition of *Escherichia coli* O157:H7 as a new enterohemorrhagenic (EHEC) pathogen, later to be determined as the causative agent of thrombotic thrombocytopenia purpura (TTP), hemorrhagic colitis (HC), and the more severe disease hemolytic-uremic syndrome (HUS). This pathogen has been implicated in numerous outbreaks in nursing homes, schools, and day care centers, typically introduced in contaminated food (beef, pork, poultry, lamb) or water. Most recent food-borne cases have been due to incompletely cooked ground beef (Swanek). Cows and chickens have been determined to be the resevoirs of this organism, and with increased epidemics in the United States, Canada, and presently in Japan, the O157:H7 pathogen is considered an important public health issue since the organism in four to eight percent of children with HC develop into HUS, with a five to ten percent fatality rate (Brian, Levine).

Clinical laboratories identify *E. coli* species though specific serological and biochemical tests. The antisera for *E.coli* do not provide information about toxins, and are not species specific as they may cross-react with other bacteria such as *Citrobacter freundii* and *Escherichia hermanii* (Cebula, Rice). O157:H7 *E. Coli* does not ferment D sorbitol as other *E. coli* species do, and MacConkey Sorbitol agar is useful to differentiate the pathogen. However, there are some sorbitol fermenting O157:H7 isolates that have been reported (Cebula). Other problems are that O157:H7 *E. Coli* organisms may cause milder symptoms than the more prominent clinical syndromes, or few pathogens may be found among the large numbers of normal microflora. Therefore quick, efficient, sensitive and

specific assays performed in minimal time and with simple equipment are needed to determine the presence of the O157:H7 pathogen.

The virulence of O157:H7 is due to the presence of a 60 megadalton (MDa) plasmid as well as the production of shigella-like toxins (SLT-I and SLT-II). A method that will identify both the plasmid and toxins (Cebula, Stacy-Phipps) would provide a more useful identification of the O157:H7 organism. Sequencing data has allowed the development of primers that identify both the plasmid and SLT's. Multiplex conditions, where both set of primers added to the same reaction, have been reported (Fratamico). It is the purpose of this work to determine if the PCR primers can be used consistently with multiplex conditions, and to be later tested with new developed techniques for quicker results.

MATERIAL AND METHODS

Specimens

E. coli and other bacterial samples were stored at - 70 C. Samples were obtained from clinical patients and animal sources, from the American Type Culture Collection (ATCC), the Center for Disease Control, the Alabama Department of Health, Auburn University, The Texas Department of Health, the Molecular Biology Institute (Scranton, Pa), and the hosting laboratory (AL/AOEL, bacteriology).

Serological and Biochemical Assays

Bacterial strains were characterized by the the VITEK GNI (bioMerieux Vitek), the Premier EHEC Enzyme Immunoassay Test (Meridian Diagnostics), *E. coli* O157 Test Kit (Diagnostic Reagents), and MacConkey Agar with and without sorbitol (Remel). Company instructions were followed for each assay.

Bacterial Lysates

Bacteria were streaked onto a brain heart infusion plate and inoculated overnight at 37 ° C. An isolated colony was inoculated into 200 ul of a solution consisting of 0.5% Triton X-100, 20 mM Tris (pH 8.0), and 2 mM EDTA, and boiled for 10 minutes.

Polymerase Chain Reaction (PCR)

The PCR reactions were performed as described by Fratamico et al. Briefly, conditions for amplification of the 60 MDa plasmid were as follows: The PCR reaction consisted of 10 mM Tris (pH 8.3), 50 mM KCL, 0.001% gelatin, 200 µM of dNTP's, 1.5 mM $MgCl_2$, 2.5 Units Taq Polymerase, and 50 pm of each primer MFSIF (5'-ACGATGTGGTTTATTCTGGA - 3') and MFSIR (5'-CTTCACGTCACCATACATAT - 3'). Five µl of bacterial lysate was added to the reaction. The mix was placed in the Perkin Elmer 9600 thermal cycler at the following conditions; an initial denaturation at 94 ° C for 5 minutes was followed by 35 cycles of 94° C for 30 seconds, 55 ° C for 30 seconds, and 72 ° C for 90 seconds, followed by extention at 72 ° C for 5 minutes. The PCR reactions were examined for the presence of a 166 bp product on an ethidium bromide stained 1.6% agarose gel in 0.5X TBE buffer. The buffer for the detection of the shigella-like toxins were the same except 2 mM $MgCl_2$ was used for the reactions. The primers that targeted both the SLT-I and SLT-II genes were: MKI 5'- TTTACGATAGACTTCTCGAC-3' and MK2 5'- CACATATAAATTATTTCGCTC - 3'. Cycling conditions were 94 ° C for 5 minutes followed by 35 cycles of 94 °C for 1 minute, 43 °C for 3 minutes, and 72 ° C for 4 minutes, followed by extention at 72 ° C for 5 minutes. PCR products of 224 bp and 227 bp were visualized on a 1.6 % agarose gel. Multiplex conditions were

performed where both primers were added to the same reaction tube. To detect both products, 2 mM $MgCl_2$ was used, and cycling conditions were: 94 ° C for 5 minutes followed by 35 cycles of 94 ° C for 1 minute , 48 ° C for 3 minutes, and 72 ° C for 4 minutes, followed by extention at 72 ° C for 5 minutes. PCR products of 166 bp and 224 bp and 227 bp were visualized on an ethidium stained 1.6% agarose gel.

Plasmid Preparation

A colony of bacterial cells was inoculated into 4 ml brain heart infusion media and incubated overnight with shaking at 37 ° C. The next day 300 µl of the overnight culture was transferred into a microcentrifuge tube, and spun 10 minutes at 7500 rpm. The supernatant was poured off, and the pellet resuspended in100 ul TE buffer and 350 µl lysis buffer (3% SDS, 50 mM Tris, pH 12.6). The tubes were inverted gently several times to mix and then incubated at 55° C for 20 minutes. After incubation, 500 µl of phenol:chloroform was added, and gently mixed by inversion for 4 minutes. The tubes were spun at 14,000 rpm for 10 minutes, and the top layer recovered and stored at 4 ° C.

Field Inversion Gel Electrophoresis (FIGE)

To separate and visualize the large molecular weight plasmids, 40 ul of plasmid preparation and 5 µl 6X running dye were loaded onto a 1% agarose gel (Seakem # 50072) in 0.5X TBE buffer. The FIGE apparatus (FIGE Mapper Electrophoresis System from Biorad) was run on program 3 (16 hours running time with the forward voltage at 180V and the Reverse Voltage at 120V). On completion of the run, the gel was stained for visualization in ethidium bromide.

RESULTS

Conventional biochemical and serological assays identified study samples as 0157:H7 *Escherichia coli*, or as another organism(s) as shown in Table I. The results were consistent with what was expected for the sample. Next, PCR reactions for the 60 MDa plasmid were performed on all samples (Table II). Again results corresponded to the clinical assays with only two exceptions; A5 and A63 which were positive by clinical assays, did not possess the plasmid by the PCR reaction. A63, interestingly, was also identified as the only positive sorbitol fermentator, and was EHEC negative suggesting this was not an O157:H7 as originally thought. These samples were retested for all assays 2 times, with the same results. The 2 samples were from patients with *E. coli*; A5 from the 1982 outbreak in Oregon and A63 from a Texas patient. The SLT-I and SLT-II genes were next tested for all samples by PCR and are shown in Table III. Results correspond to what is expected; *E. coli* 0157:H7 samples often possess the SLT-I or SLT-II toxins, while control strains, such as different Shigella species or laboratory strains are negative for these genes.

The multiplex conditions were tested to determine whether the plasmid, the SLT-I, and the SLT-II genes could be simultaneouly amplified. Preliminary results are shown in figure 1; negative samples, DH5 α, H_2O, and lysate showed no bands. A5, A13, and A14, postive for only the SLT, showed 1 band at 224-227 bp (agarose gels are unable to distinquish between 224 bp and 227 bp; the differentiation between the two toxins will be tested by the Bio-Rad Universal Gene Comb Kit, and written by Leigh K. Hawkins). A57, as expected, showed only 1 band at the 166 bp position, while A45, positive for both genes, showed bands at the 166 bp and 224 bp positions. Although non-specific bands appeared with the lower stingency PCR conditions, they do not interfere in the Gene Comb assay (unpublished observations). The multiplex assay was performed on a number of samples for reliability in the detection of both plasmid and SLT genes, and Table IV shows the results. The multiplex conditions amplified both

genes and corresponded with high reliability with the results of each gene when tested separately. The amplified products were then tested on the Bio-Rad Gene Combs to differentiate between the plasmid gene, the SLT-I gene, and the SLT-II gene (Leigh K. Hawkins).

The reason why A5 and A63 are atypical was tested by observing their plasmids by FIGE. Figure 2 shows the results; A47, an ATCC *E. coli* wildtype strain and A45, a positive strain show the bands for the plasmid, while DH5 and A39, the negative controls, do not show any plasmid bands. Samples that are *E. coli* species from different sources, but not O157:H7 , are A28, A42, and A46. The bands of A28 have higher molecular weight bands than the controls, while A42 does not appear to have a plasmid. A46 bands are similar to A47 and A45, but are slightly lower in molecular weight; since many *E. coli* species do have plasmids, this was not unexpected and the PCR results confirmed that the 60 MDa plasmid was not present. Interesting were the Texas Department of Health (TDH) samples. A62, A78, and 98 had similar plasmids, and are the same molecular weight as the wildtypes, A47 and A45. They also had the same results in the biochemical, plasmid and SLT PCR, and multiplex PCR reactions. A11 had similar results to the other *E. coli* O157:H7; however the FIGE shows 4 plasmid bands. A63 is not an O157:H7 since it is sorbital positive, EHEC negative, plasmid negative, and SLT negative. The FIGE gel shows a plasmid that has many bands and which is very different when compared to the other plasmids (Figure 2). A5 which is O157:H7, but negative in the plamid PCR reaction does not have a plasmid except at very low molecular weight. Currently A5, A11, and A63 plasmid bands are under study.

Discussion

O157:H7 virulence is in part attributed to the presence of the 60 mDa plasmid and toxigenic genes, SLT-I or SLT-II. The ability to assay specifically for these genes would provide a quick, distinguishing diagnosis of the *E. coli* O157:H7 pathogens. These tests must be specific, sensitive, and with a minimal time requirement if they are to replace conventional biochemical assays. The multiplex assay described by Fratamico was found by the study to meet the requirements of specificity and sensitivity, and presently are being evaluated in conjunction with the Bio-Rad Gene Comb Test Kit to determine whether this assay would full-fill the criteria. However, as seen from the results and confirmed by the FIGE assay, some O157:H7 *E. coli* have variants which must be tested for and not overlooked (Brian, Rice). The addition of other sets of primers specific for these organisms, the uidA gene, the eaeA gene and the hemolysin gene, could improve the recognition of these variants (Fratamico, Cebula, & Karch). These primers are presently being considered in the laboratory.

The atypical sample A5 is thought to be due to the difference in the plasmid genes. Currently the laboratory in collaboration with Scranton University, is sequencing the plasmid, and once completed, comparison of atypical strains to the wild type may be determined through sequencing or hybridizations. Testing for other atypical strains and research on their plasmids will be continued in the laboratory.

Therefore it may be concluded that multiplex PCR is a reliable technique for the detection of enterohemorrhagic E. coli O157:H7. However additional assays may be required for detection of the variants.

ID #	O157:H7	O157 Ag (agglutination)	Sorbitol (Vitex)	EHEC
A5	Y	+	-	+
A6	Y	+	-	
A7	Y	+	-	
A8	Y	+	-	
A9	Y	+	-	
A10	Y	+	-	
A11	Y	+	-	+
A12	Y	+	-	
A13	N	-	+	-
A14	N	-	+	+
A15	Y	+	-	
A16	Y		-	
A17	Y		-	
A18	Y		-	
A19	Y		-	
A20	Y		-	
A21	Y	+	-	+
A23	N	-	+	
A24	N	-	+	-
A25	N	-	+	
A28	N	-	+	-
A39	?	-	+	
A40	N	-	+	-
A41	N	-	+	-
A42	N	-	+	-
A43	N	-	+	-
A44	N	-	+	-
A45	Y	+	-	+
A46	N	+	+	-
A47	?	+	-	
A48	N	-	+	-
A49	Y	+	-	+
A50	Y	+	-	+
A51	N	-	+	-
A57	Y	+	-	-
A58	Y	+	-	+
A59	Y	+	-	+
A60	Y	+	-	+
A61	Y	+	-	+
A62	Y	+	-	+
A63	Y	+	+	-
A64	Y	+	-	+
A65	Y	+	-	+
A66	Y	+	-	+
A67	Y	+	-	+
A68	Y	+	-	+

A69	Y	+		+
A70	Y	+	-	+
A71	Y	+	-	+
A72	Y	+	-	
A73	Y	+	-	
A74	Y	+	-	
A75	Y	+	-	
A76	Y	+	-	
A77	Y	+	-	
A78	Y	+	-	+
A79	Y	+	-	+
A80	Y	+	-	
A81	Y	+	-	
A82	Y	+	-	
A83	Y	+	-	
A84	Y	+	-	
A85	Y	+	-	
A86	Y	+	-	+
A87	Y	+	-	
A88	Y	+	-	
A89	Y	+	-	
A90	Y	+	-	
A91	Y	+	-	
A92	Y	+	-	
A93	Y	+	-	
A94	Y	+	-	
A95	Y	+	-	
A96	Y	+	-	
A97	Y	+	-	
A98	Y	+	-	+

TABLE II

PCR RESULTS

MS1F / MS1R

	166 bp fragment
pos control	A45
neg control	DH5
A5	negative
A6	positive
A7	positive
A8	positive
A9	positive
A10	positive
A11	positive
A12	positive
A13	negative
A14	negative
A15	positive
A16	positive
A20	positive
A21	positive
A23	negative
A24	negative
A39	negative
A40	negative
A41	negative
A47	positive
A48	negative
A49	positive
A50	positive
A51	negative
A57	positive
A58	positive
A59	negative
A60	positive
A61	positive
A62	positive
A63	negative
A64	positive
A66	positive
A67	positive
A68	positive
A82	positive
A112	positive
A124	positive
A125	negative
A126	negative

TABLE III

PCR RESULTS

	MK1 / MK2 224bp / 227bp
pos control	A 45 , A13, & A14
neg control	DH5
A5	positive
A6	NT
A7	NT
A8	negative
A9	NT
A10	NT
A11	positive
A12	NT
A13	positive
A14	positive
A15	NT
A16	NT
A20	NT
A21	NT
A23	NT
A24	NT
A39	NT
A40	NT
A41	NT
A47	positive
A48	NT
A49	positive
A50	positive
A51	NT
A57	negative
A58	positive
A59	NT
A60	NT
A61	NT
A62	positive
A63	negative
A64	positive
A66	positive
A67	NT
A68	NT
A82	NT
A112	NT
A124	NT
A125	NT
A126	NT

TABLE II

A127	negative	A127	NT
A128	negative	A128	negative
A129	negative	A129	negative
A130	negative	A130	NT
A131	negative	A131	NT
A132	negative	A132	NT
A133	negative	A133	NT
A134	negative	A134	NT
A135	positive	A135	negative
A136	negative	A136	negative
A137	negative	A137	NT
A139	negative	A139	NT
A140	negative	A140	NT
A141	negative	A141	negative
A142	negative	A142	negative
A146	negative	A146	NT
A150	negative	A150	NT
A151	negative	A151	NT
A152	negative	A152	negative

SAMPLE	COMMENT	Plasmid	SLT	pO157	pO157
				miniscreen	other plasmid
A 5	plasmid negative & SLT positive from 1982 Oregon outbreak	negative	positive	-	40 kb
A 11	plasmid negative & SLT positive from 1992 Maine patient	positive	positive		
A 21	isolate from a cow	positive	positive		
A 28	environmental isolate from P.R. not a O157	negative	negative		
A 39	XL1 blue cells (Stratagene) negative control	negative	negative	-	
A 42	patient with urinary infection not O157	negative	negative		
A 45	plasmid positive & SLT positive positive control	positive	positive	+	4 - 6 kb
A 46	E. coli from hamburger not O157	negative	negative		
A 47	ATCC # 43894 wild type E. coli	positive	positive	+	
A 48	DH1pSa = lab strain with 40 Kb	negative	negative	-	
A 52	4.0 kb (=pZdeltaB23C373) = used to bump pF-like large plasmid O157:H7 strains	negative	negative		
A 57	ATCC #43888 plasmid, no SLT I or II genes	positive	negative		
A 59	w.t. plus a 4 kb plasmid (minus the large plasmid)	negative	positive		
A 62	Patient from Texas Department of Health	positive	positive	+	
A 63	Texas Department of Health patient that is plasmid negative & SLT negative	negative	negative	+	muliple
A 78	Patient from Texas Department of Health	positive	positive	+	
A 98	Patient from Texas Department of Health	positive	positive	+	
DH5	negative control	negative	negative	-	

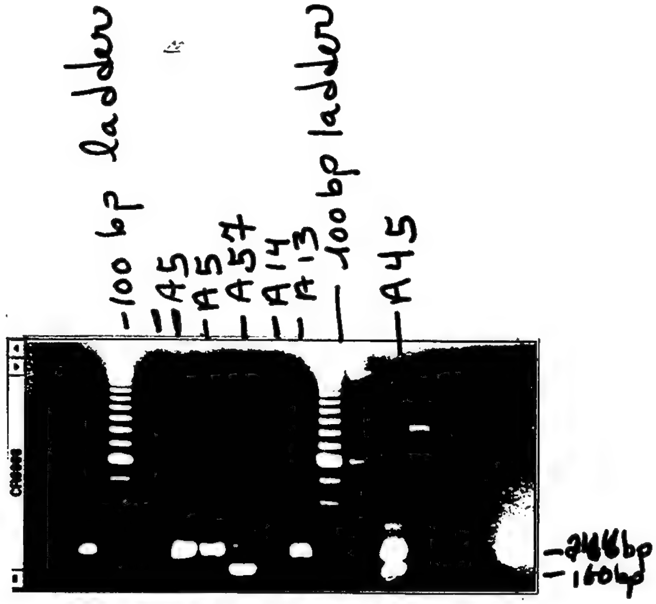


Figure 1. Multiplex PCR Reactions

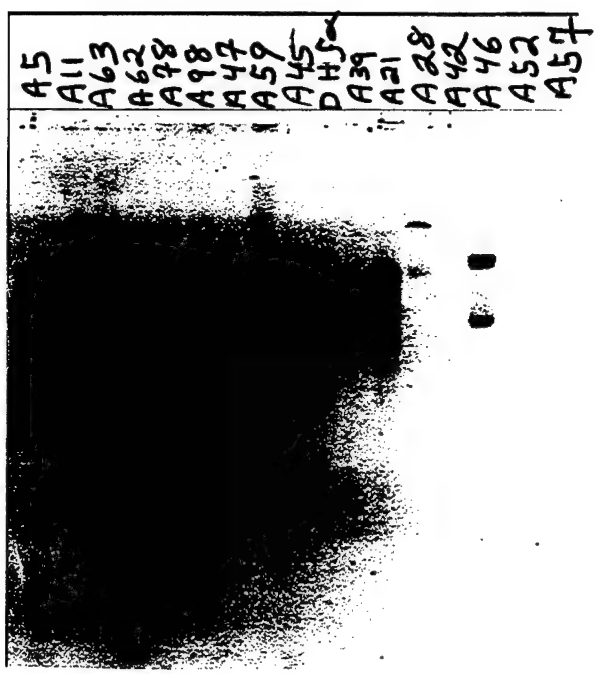


Figure 2. FIGE results on large plasmid preparations

THE EFFECT OF SHORT DURATION RESPIRATORY MUSCULATURE TRAINING ON TACTICAL AIR COMBAT MANEUVER ENDURANCE AND RECOVERY

Michael E. Rogers
Doctoral Candidate
School of Exercise, Leisure and Sport

Kent State University
163 MACC Annex
Kent, OH 44242-0001

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

August 1996

THE EFFECT OF SHORT DURATION RESPIRATORY MUSCULATURE TRAINING ON TACTICAL AIR COMBAT MANEUVER ENDURANCE AND RECOVERY

Michael E. Rogers
Doctoral Candidate
School of Exercise, Leisure and Sport
Kent State University

Abstract

The purpose of this study is to test the effect of a low frequency, short duration respiratory musculature training (RMT) program on tactical air combat maneuver (TACM) endurance in experienced acceleration research subjects. Also of interest is the effect of RMT on recovery from the TACM. While studies such as these have been conducted examining general strength training of the skeletal musculature, only minimal acceleration research has been conducted utilizing specific training of the respiratory musculature, despite the critical role the respiratory musculature plays in the intra-thoracic pressure component of the anti-G straining maneuver. Subjects will be randomly assigned to either the respiratory musculature training group (RMT) or the control group. Variables [units] to be measured include: peak heart rate response to TACM [beats·min^{-1}], TACM endurance [s], subject perceived effort (ordinal scale) during TACM, subject recovery time from TACM [s difference between TACM replications], aerobic capacity (VO_2 peak) [ml·kg^{-1}·min^{-1}], peak and mean anaerobic power [W], miscellaneous body composition variables (fat mass, fat-free mass, etc.), maximal inspiratory/expiratory pressures [mm Hg], maximal inspiratory/expiratory capacities [ml], maximal inspiratory/expiratory flow rates [ml·s^{-1}]. My role during the 10-week tour was to revise the original proposal for review board approval and to organize the laboratory so that they may initiate data collection on this project.

THE EFFECT OF SHORT DURATION RESPIRATORY MUSCULATURE TRAINING ON TACTICAL AIR COMBAT MANEUVER ENDURANCE AND RECOVERY

Michael E. Rogers

I. Introduction

The anti-G straining maneuver (AGSM) currently taught to USAF aircrew of high performance aircraft has two principal components (1). The first, a skeletal muscular component, requires the sustained, isometric recruitment of the skeletal musculature. Second, the intra-thoracic pressure component of the AGSM involves repeated 3-s cycles of forceful expiratory efforts against a closed or partially closed glottis (1). Repetition of an effective AGSM is of critical importance to fighter aircrew in the sustained high $+G_Z$ environment of air combat maneuvering to aid venous return and maintain eye-level arterial blood pressure.

Because of the extreme physical demands placed on individuals during repeated performance of the AGSM during high $+G_Z$ conditions, one might assume that physically untrained individuals will fatigue sooner than their physically trained peers (2). Addressing the skeletal musculature component of the AGSM, research has been conducted to confirm and quantify the protective effect of general strength training programs (i.e., whole body resistance training) on $+G_Z$ endurance (3-7). While studies such as these have been conducted examining general strength training of the skeletal musculature, only minimal acceleration research has been conducted utilizing specific training of the respiratory musculature, despite the critical role the respiratory musculature plays in the intra-thoracic pressure component of the AGSM. The research which has been conducted in acceleration utilizing respiratory musculature training (RMT) has demonstrated that such training can result in an improved AGSM (8,9) but has not been specifically applied to a simulated operational $+G_Z$ endurance profile. It has been demonstrated that increased respiratory work during the AGSM results in respiratory muscle fatigue and this fatigue correlates with early termination of centrifuge exposures (10). In addition, it is possible to elicit fatigue in the diaphragm (11) and other respiratory muscles (12), and it has been demonstrated that appropriate training increases both strength and endurance of the

respiratory musculature (13). However, the strength/endurance training regimens detailed in the literature are time consuming (long duration) and not practical for aircrew use (13).

The purpose of this study is to test the effect of a low frequency, short duration RMT program on the TACM endurance in experienced acceleration research subjects. Also of interest is the effect of RMT on recovery from the TACM. This study employs the specificity of exercise training principle (14) relative to the intra-thoracic pressure component of the AGSM.

II. Objective

The purpose of this experiment is to evaluate the effectiveness of strength/endurance training of the respiratory musculature (short duration) in enhancing tactical air combat maneuver (TACM) endurance and recovery.

III. Impact Statement

This experiment is an important step in responding to the flying community's request for a training program which enhances $+G_Z$ endurance. If this study is not conducted, a unique opportunity to effectively and efficiently enhance $+G_Z$ endurance may be lost.

IV. Methodology

A. Equipment and Facilities:

1. Centrifuge: Subjects' TACM endurance will be assessed from standard measurement equipment during exposures on the Armstrong Laboratory Centrifuge. Subjects will be experienced members of the acceleration subject panel and will be fitted with the Advanced Technology Anti-G Suit (ATAGS), HGU-55/P helmet and MBU-20/P oxygen mask for all centrifuge runs with the exception of a gradual onset run (GOR) to establish relaxed $+G_Z$ tolerance, where no helmet/mask configuration will be worn. To minimize breathing resistance from the oxygen mask during the AGSM, a high-flow CRU-94 oxygen regulator, in

the "emergency" setting, will be used for all centrifuge exposures requiring subjects to wear a mask. All subjects will be fitted with a 5-lead ECG for two-axis recording of heart rhythm and heart rate during all centrifuge exposures. Analog data signals of physiologic and anti-G suit pressure variables will be recorded on a multiple channel strip chart recorder (Gould 2800S) with magnetic tape back-up for subsequent data reduction and statistical analysis. Centrifuge subjects will be observed via closed circuit video monitors in the centrifuge control room.

2. Ergometry: Ergometry tests will be conducted in the Applied Exercise and Work Physiology Laboratory (AL/CFT) in Building 170. The Applied Exercise and Work Physiology Laboratory (AL/CFT) conducts all exercise testing according to guidelines and recommendations set forth by the American College of Sports Medicine (15).
   a. Aerobic Capacity: A motorized treadmill (Quinton Q-65) will be used to elicit work for the aerobic capacity test. Expired gas will be collected from subjects during tests and analyzed utilizing a Rayfield Gas Collection System equipped with Ametek O_2/CO_2 analyzers. Blood pressure will be monitored using a Finapres finger cuff and monitor. Subjects will also be fitted with a 3-lead ECG configuration for monitoring of heart rate and rhythm using a Quinton Q710 electrocardiograph.
   b. Anaerobic Power: Anaerobic power will be determined from administration of a 30-s maximal pedaling test on a cycle ergometer (Cardionics).

3. Body Composition: Body composition measurements will be conducted in the Auxiliary Exercise Physiology Laboratory in Building 170.
   a. Body Density: Body density will be estimated via hydrostatic weighing using a 300-gallon hydrostatic chamber equipped with a 15 kg X 20 g autopsy scale (Chatillon). Residual volume will be determined on land prior to submersion via oxygen dilution using O_2 and CO_2 analyzers (Ametek).
   b. Total Body Water: Total body water/extracellular water will be estimated utilizing a bioelectrical impedance analysis system (Xitron 4000B).

c. Anthropometry: A series of anthropometrical measurements will be made on subjects to aid in anti-G suit fitting and to serve as a secondary evaluation of body composition. Measurements include skinfold (Lange calipers), body circumference (Lufkin tape), bone breadths (Lafayette calipers), body mass (Detecto scales) and body stature (Detecto measuring rod).

4. Respiratory Variables: Respiratory tests will be conducted in the Applied Exercise and Work Physiology Laboratory (AL/CFT) and Auxiliary Exercise Physiology Laboratory in Building 170.

   a. Respiratory Pressures: An electronic pressure manometer (Omega) will be used to measure maximal expiratory pressure and maximal inspiratory pressure of all subjects.
   b. Lung Capacities: A Collins 120-L spirometer will be used to measure lung capacities.
   c. Pulmonary Flow Rates: A Collins 13.5-L spirometer will be used to measure pulmonary flow rates.

B. Subjects:

Both male and female subjects from the Armstrong Laboratory acceleration subject panel who have maintained acceleration training standards will be recruited for this study. We will recruit sufficient subjects to have approximately 20 subjects (10 treatment, 10 control) complete the study. To qualify for the study, subjects must have recent $+G_Z$ exposure experience and be able to complete at least four consecutive cycles (+52 s) of the TACM while wearing ATAGS. One TACM cycle consists of 5 s at $+9\ G_Z$, 1 s at $+5\ G_Z$, 5s at $+8\ G_Z$ and 2 s at $+4\ G_Z$. Acceleration panel subjects who have not routinely performed TACM exposures with helmet and mask will perform TACM familiarization exposures wearing the full equipment ensemble, including helmet and mask, prior to establishment of maximal baseline TACM endurance times. Panel subjects also will receive RMT program indoctrination (experimental group only) before they perform respiratory testing and initiate their 6-week RMT regimen. Only acceleration panel subjects who have no history of prolonged resistance training participation and have not participated in muscular strength training programs during the past four weeks will be studied. To control for

the potentially detrimental effects of excessive aerobic training on +Gz tolerance (3), subjects who perform regular aerobic activities equal or equivalent to running more than 15 miles per week also will not be included in the study. In addition, only non-smoking subjects will be recruited.

C. Duration of the Study:

Determination of relaxed $+G_Z$ tolerance, acceleration variables, respiratory variables, ergometric variables and body composition variables will require approximately two weeks. These determinations will be performed prior to (pre) and following (post) the 6-week RMT period. Duration of the entire study is dependent upon the timeliness of subject recruitment. However, depending on the time required to complete pre- and post-RMT testing, duration of participation for each individual subject will be approximately 10 weeks. Every attempt will be made to have subjects adhere to the following 10 week testing/training schedule:

a. Week 1, Monday: indoctrination to respiratory testing and RMT (treatment group only) procedures
b. Week 1, Tuesday: centrifuge testing for $+G_Z$ tolerance (pre-RMT 1)
c. Week 1, Wednesday: aerobic capacity testing (pre-RMT)
d. Week 1, Thursday: body composition testing (pre-RMT)
e. Week 1, Friday: anaerobic power testing (pre-RMT)
f. Week 2, Monday: indoctrination to respiratory testing and RMT (treatment group only) procedures
g. Week 2, Tuesday: centrifuge testing for $+G_Z$ tolerance (pre-RMT 2)
h. Week 2, Wednesday: respiratory testing (pre-RMT)
i. Weeks 3-8: RMT (5 days per week) for treatment group only and centrifuge $+G_Z$ proficiency maintenance (1 day per week) for all subjects (i.e., RMT and control groups)
j. Week 9, Monday: re-acquaintance with respiratory testing procedures
k. Week 9, Tuesday: centrifuge testing for $+G_Z$ tolerance (post-RMT 1)
l. Week 9, Wednesday: aerobic capacity testing (post-RMT)
m. Week 9, Thursday: body composition testing (post-RMT)

n. Week 9, Friday: anaerobic power testing (post-RMT)

o. Week 10, Monday: re-acquaintance with respiratory testing procedures

p. Week 10, Tuesday: centrifuge testing for $+G_Z$ tolerance (post-RMT 2)

q. Week 10, Wednesday: respiratory testing (post-RMT)

D. Experimental Procedures:

1. General Experimental Design: The general experimental design entails a single between-subjects effect and no within-subject effects. There are two levels of the between-subjects effect (i.e., RMT, control). Subjects will be randomly assigned to either the RMT or the control group. Variables [units] to be measured include:

a. peak heart rate response to TACM [beats$\cdot$min^{-1}]

b. TACM endurance [s]

c. subject perceived effort (ordinal scale) during TACM

d. subject recovery time from TACM [s difference between TACM replications]

e. aerobic capacity (VO_2 peak) [ml $\cdot$kg$^{-1}\cdot$min^{-1}]

f. peak and mean anaerobic power [W]

g. miscellaneous body composition variables (fat mass, fat-free mass, etc.)

h. maximal inspiratory/expiratory pressures [mm Hg]

i. maximal inspiratory/expiratory capacities [ml]

j. maximal inspiratory/expiratory flow rates [ml$\cdot$s^{-1}]

2. Centrifuge Procedures:

a. Testing for $+G_Z$ Tolerance: $+G_Z$ tolerance will be tested for pre- and post-RMT training. Although peak relaxed $+G_Z$ tolerance will be established and periodically re-assessed for all subjects, the principal criterion by which effectiveness of the exercise conditioning regimens will be evaluated is the time (s) subjects can endure the TACM profiles. For all centrifuge runs, the centrifuge will be configured with the upright (13° seat back angle) ACES II-type seat. Foot plates in the centrifuge gondola will be positioned to provide an approximately

120° knee joint angle. Each subject will be properly fitted with ATAGS and a HGU-55/P helmet with MBU-20/P mask for all centrifuge exposures with the exception of the GOR to establish relaxed $+G_Z$ tolerance where no helmet/mask will be worn. The pre- and post-RMT centrifuge testing will be conducted in duplicate. Acceleration profiles for pre- and post-RMT data collection consist of:

i. GOR ($0.1\ +G_Z \cdot s^{-1}$) to peak $+G_Z$ or other endpoint criteria (medical monitor decision, self-initiated decision, nausea, 100% peripheral light loss or 50% central light loss*), relaxed (no AGSM), anti-G suit worn but NOT pressurized

ii. Warm-up consisting of 1 cycle of ROR ($6.0\ +G_Z \cdot s^{-1}$) to $+5.0\ G_Z$ for 15 s, with AGSM, with anti-G suit pressurized

iii. 2-min rest period

iv. TACM to fatigue or other end-point criteria (medical monitor decision, self-initiated decision, nausea, 100% peripheral light loss or 50% central light loss*), with AGSM, with anti-G suit pressurized, no central observer coaching

v. 5-min rest period

vi. TACM to fatigue or other end-point criteria (medical monitor decision, self-initiated decision, nausea, 100% peripheral light loss or 50% central light loss*), with AGSM, with anti-G suit pressurized, no central observer coaching

*if $+G_Z$ loss of consciousness (G-LOC) occurs it will constitute termination of $+G_Z$ endurance testing and, consequently, time to G-LOC (s) will be used as the endurance variable.

b. $+G_Z$ Proficiency Maintenance: In order to maintain their entry level AGSM proficiency and $+G_z$ endurance baseline, subjects will be required to perform a TACM centrifuge training profile once a week throughout the RMT period. This $+G_z$ training requirement consists of a warm-up chosen by the subject, followed by a submaximal TACM profile (i.e., ≈75% of the pre-RMT TACM performance) during first four weeks of RMT. During weeks 5 and 6 of the RMT, subjects will be allowed to perform maximally during their weekly centrifuge "maintenance" runs (i.e., subjects will be allowed to perform the TACM to fatigue). Those subjects in the control group, who will not be engaging in RMT, will be

required to follow the same centrifuge maintenance schedule as those subjects participating in RMT. Any subjects missing more than one of their weekly centrifuge maintenance rides over the 6-week training period will be removed from the study.

3. Ergometry Procedures:

Tests for aerobic capacity and anaerobic power will be conducted pre- and post-RMT training.

a. Testing for Aerobic Capacity: Standard electrocardiographic electrodes will be applied prior to performing the test for aerobic capacity so heart rhythm and rate can be monitored. Blood pressure will be monitored throughout the test for aerobic capacity via oscillometry using the Finapres monitoring system. Expiratory gasses will be collected and analyzed using the Rayfield Gas Analysis System. The test will continue to volitional or physiological maximal effort, which ever occurs first. Specific reasons for test termination include (15,16):

    i. subject request to terminate test
    ii. subject is unable to continue test (i.e., volitional maximum)
    iii. failure of subject oxygen consumption to increase with an increase in work demand with confirmation from elevated respiratory exchange ratio (>1.15) and/or heart rate greater than age estimated maximum
    iv. onset of angina or angina-like symptoms
    v. significant drop (20 mm Hg) in systolic blood pressure or failure of systolic blood pressure to increase in response to an increase in work demand
    vi. excessive rise in blood pressure (systolic>260 mm Hg; diastolic>115 mm Hg)
    vii. subject exhibits signs of poor perfusion such as lightheadedness, confusion, ataxia, pallor, cyanosis, nausea or cold/clammy skin
    viii. ECG abnormalities such as a pronounced ECG change from baseline (i.e., > 2mm of horizontal/down sloping ST-segment depression or >2mm of ST-segment elevation), less serious arrhythmias (supraventricular tachycardia) and exercise-induced bundle branch block indistinguishable from ventricular tachycardia
    ix. dyspnea
    x. equipment failure

The treadmill protocol for aerobic capacity is that derived by Ellestad (17). This protocol requires subjects to exercise continuously to VO_{2peak} according to the following grade/belt rate progression:

Stage	Minutes	Grade (%)	Belt rate ($mile \cdot h^{-1}$)
1	0:00- 3:00	10.0	1.7
2	3:01- 5:00	10.0	3.0
3	5:01- 7:00	10.0	4.0
4	7:01-10:00	10.0	5.0
5	10:01-12:00	15.0	5.0
6	12:01-14:00	15.0	6.0
7	14:01-16:00	15.0	7.0
8	16:01-18:00	15.0	8.0

b. Testing for Anaerobic Power: The Wingate Test (18,19) will be used to determine peak and mean anaerobic power. The test is a 30-s maximal cycle ergometer test. The warm-up for the test consists of 1) 1 min of pedaling (50 rpm) against no resistance; 2) 3 min of pedaling (50 rpm) against a moderate load (50% of actual test resistance) interspersed with three sprints (maximal rpm) of 5 s each and 3) 1 min of pedaling (50 rpm) against no resistance. After a brief rest period (3-5 min), the subject pedals as fast as possible against the test workload ($0.09\ kg \cdot kg$ body $mass^{-1}$) for 30 s. Recovery includes pedaling against a minimal resistance (≈ 1.0 kg) for at least 3-5 min. Premature termination of the test can occur for the following reasons:

i. subject request to terminate test
ii. subject is unable to continue test
iii. onset of angina or angina-like symptoms
iv. subject exhibits signs of poor perfusion such as lightheadedness, confusion, ataxia, pallor, cyanosis, nausea or cold/clammy skin
v. dyspnea
vi. equipment failure

4. Body Composition Procedures:

Body composition measurements will be made pre- and post-RMT training. Subjects will be asked to report for testing in a post-absorptive state following an overnight fast. Subjects will be asked to maintain normal hydration during this fasting period.

a. Body Density: Body volume will be determined via hydrostatic weighing. From this body volume estimate body density will be calculated for use in a multi-component model with total body water to estimate total body adiposity. Subjects will be weighed underwater while forcefully expiring to residual lung volume. The procedure will be repeated until successive attempts fail to produce an increase in underwater weight. Residual lung volume will be measured on land prior to the hydrostatic weighing procedure. The modified oxygen dilution method (20) for measuring residual lung volume will be used. This technique requires subjects to forcefully expire to residual volume, hyperventilate 100% O_2 and forcefully expire to residual volume again. The gas collected after the final expiration will be analyzed using Ametek O_2 and CO_2 analyzers.

b. Total Body Water: Bioelectrical impedance analysis will be used to estimate total body water content. Total body water will be used in a multi-component model with body density to estimate total body adiposity. The procedure requires subjects to lie supine and have four superficial electrodes placed on the right side of his/her body at the following anatomical sites:

   i. posterior aspect of the hand over the distal portion of the metacarpals
   ii. posterior aspect of the arm just superior to the carpals
   iii. superior aspect of the foot over the distal portion of the metatarsals
   iv. anterior aspect of the leg just superior to the tarsals

A measurement current of 500μA is passed between the electrodes for the determination of the impedance to this electrical flow provided by the circuit (i.e., the subject's body). This device is listed with the U.S. Food and Drug Association as a Class II Medical Device and has been designed to meet the safe current limits standards for patient connected equipment. The impedance data will be reduced using the Xitron 4000B with computer interface.

c. Anthropometry: Subjects will undergo a series of anthropometrical measurements for body composition estimation (21, 22) and to assist in anti-G suit fit. The anthropometrical measurements include skinfolds, body circumferences, bone breadths, body mass and stature.

5. Procedures for Respiratory Musculature Testing:

Measurements of inspiratory and expiratory maximal pressures, capacities and flow rates will be made pre- and post-RMT training. Testing will include:

a. Maximal Inspiratory Pressure: The determination of maximal inspiratory pressure requires subjects (seated with nose clip applied) to first expire maximally to residual lung volume through a plastic cylinder connected to a pressure manometer (Omega) which will measure negative pressure during the subsequent inspiration. When maximal expiration has occurred the subject will inspire maximally. Shortly after inspiration begins ($\approx$0.5 s), the cylinder will be occluded requiring the subject to inspire against a closed circuit. The inspiration will last approximately 3 s and a reading of negative pressure made. Subjects will be instructed how to perform the task to ensure minimal pressure production by the facial muscles. The test will be conducted in triplicate.

b. Maximal Expiratory Pressure: The determination of maximal expiratory pressure requires subjects (seated with nose clip applied) to first inspire maximally to total lung capacity through a plastic cylinder connected to a pressure manometer (Omega) which will measure positive pressure during the subsequent expiration. When maximal inspiration has occurred the subject will expire maximally. Shortly after expiration begins ($\approx$0.5 s), the cylinder will be occluded requiring the subject to expire against a closed circuit. The expiration will last approximately 3 s and a reading of positive pressure made. Subjects will be instructed how to perform the task to ensure minimal pressure production by the facial muscles. The test will be conducted in triplicate.

c. Inspiratory Capacity: The determination of inspiratory capacity requires subjects (seated with nose clip applied), at the end of resting expiration, to inspire maximally to total lung capacity through a plastic cylinder connected to a spirometer (Collins) to measure gas volume. The test will be conducted in triplicate.

d. Expiratory Capacity: The determination of expiratory capacity requires subjects (seated with nose clip applied), at the end of resting inspiration, to expire maximally to residual lung volume through a plastic cylinder connected to a spirometer (Collins) to measure gas volume. The test will be conducted in triplicate.

e. Maximal Inspiratory Flow Rate: The determination of maximal inspiratory flow rate requires subjects (seated with nose clip applied) to first expire to residual volume and then inspire maximally through a plastic cylinder connected to a spirometer (Collins). Flow/volume curves will be recorded such that flow rates can be calculated at various points throughout the inspiration. The test will be conducted in triplicate.

f. Maximal Expiratory Flow Rate: The determination of maximal expiratory flow rate requires subjects (seated with nose clip applied) to first inspire to total lung volume and then expire maximally through a plastic cylinder connected to a spirometer (Collins). Flow/volume curves will be recorded such that flow rates can be calculated at various points throughout the expiration. The test will be conducted in triplicate.

g. Maximal Voluntary Ventilation (Maximal Breathing Capacity): The determination of maximal voluntary ventilation requires subjects (seated with nose clip applied) to breathe as deeply and rapidly as possible for 15 s through a plastic cylinder connected to a spirometer (Collins) to measure gas volume. The test will be conducted in triplicate.

6. Procedures for Respiratory Musculature Training:

RMT will be conducted over a 6-week period. RMT subjects will train five days per 6-day week (i.e., training sessions can be scheduled Monday-Saturday). Within the 6-week RMT regimen, subjects will be allowed to miss a total of six training sessions. If the subjects miss more than six training sessions or miss consecutive training sessions, they will be removed from the study. A training supervisor will indoctrinate all subjects in the treatment group (i.e., not control subjects) on the RMT tasks. All RMT will require subjects to be seated in an ACES II upright seat (13° seatback angle), located in the Auxiliary Exercise Physiology Laboratory in Building 170, which is similar to the one used in the centrifuge gondola. The foot plates on this mock-up will be situated so that the subject knee joint angle is at approximately 120°. Each training session will consist of:

a. Respiratory Strength Training: RMT subjects (seated with nose clip applied) will be asked to inspire to total lung capacity with an open glottis and maintain for 10 s. Subjects will then expire maximally to residual volume and maintain for 10 s. The respiratory cycle will be repeated three times in week 1 and week 2 of the RMT program. To provide an overload, the respiratory cycle will be repeated four times in weeks 3 and 4 and will be repeated five times in the final two weeks of the RMT program. Time to complete this portion of the RMT session will be approximately 3-5 min.

b. Resistance Endurance Training: RMT subjects (seated with nose clip applied) will be asked to inspire/expire through a mouthpiece connected to a plastic tube. The tube will be narrowed to provide resistance to air flow. The stenosis will be such to provide a resistance which results in an inspiratory pressure equal to approximately 70% of the pre-RMT maximal inspiratory pressure. To provide an overload, the duration of this resistive respiration bout will be 8 min in week 1 of the RMT program and will increase by 1 min each week after that (i.e., on week 6 of the RMT program the duration will be 13 min). Time required to complete this portion of the RMT session will range from 10-15 min.

7. Statistical Analysis:

A sample consisting of 20 subjects (RMT group n_1=10 and control group n_2=10) is desired to enable the reasonable likelihood of detecting truly significant effects. An analysis of variance (1 treatment with 2 levels) will be utilized where the subjects' aerobic and anaerobic capacities will be introduced as covariates into the general linear model. The level of significance will be set a priori at $p<0.05$.

References

1. Gillingham KK, Fosdick JP. High-G training for fighter aircrew. Aviat. Space Environ. Med. 1988; 58: 12-9.
2. Gillingham KK. High-G stress and orientational stress: physiological effects of aerial maneuvering. Aviat. Space Environ. Med. 1988; 59(11, Suppl.): A10-A20.
3. Burton RR. Simulated aerial combat maneuvering tolerance and physical conditioning: current status. Aviat. Space Environ. Med. 1986; 57: 712-714.
4. Epperson WL, Burton RR, Bernauer EM. The influence of differential physical conditioning regimens on simulated aerial combat maneuvering tolerance. Aviat. Space Environ. Med. 1982; 53(11): 1091-1097.
5. Epperson WL, Burton RR, Bernauer EM. The effectiveness of specific weight training regimens on simulated aerial combat maneuvering G-tolerance. Aviat. Space Envrion. Med. 1985; 56: 534-539.
6. Mizumoto C, Iwane M. The effects of isotonic training on $+G_Z$ tolerance. Japanese J. Aerospace Environ. Med. 1984; 25: 15-29.
7. Tesch PA, Hjort H, Balldin UI. Effects of strength training on G tolerance. Aviat. Space Environ. Med. 1983; 54(8): 691-695.
8. Geng X. The strengthening of respiratory muscles and the $+G_Z$ tolerance. Chinese J. Aviat. Med. 1991; 2(3): 133-136.
9. Wang S, Zhing J, Geng X. Enhancement of anti-G effect of L-1 maneuvering by strengthening of respiratory muscles in pilots. Chinese J. Aviat. Med. 1993; 4(2): 90-92.
10. Bain B, Jacobs I, Buick F. Respiratory muscle fatigue during simulated air combat maneuvering (Abstract). Aviat. Space Environ. Med. 1995; 66(5): 470 (#41).
11. Roussos CS, Macklem PT.. Diaphragmatic fatigue in man. J. Appl. Physiol. 1977; 43: 189-197.
12. Roussos CS, Fixley M, Gross D Macklem PT. Fatigue of the inspiratory muscles and synergic behavior. J. Appl. Physiol. 1979; 46: 897-904.
13. Leith DE, Bradley M. Ventilatory muscle strength and endurance training. J. Appl. Physiol. 1976; 41: 508-516.

14. Sale D, MacDougall D. Specificity in strength training: a review for the coach and athlete. Science Periodical on Research and Technology in Sport. Ottawa, The Coaching Association of Canada, March 1981.

15. American College of Sports Medicine. ACSM's Guidelines for Exercise Testing and Prescription (5th Edition). Baltimore, Williams and Wilkins. 1995: 77-78.

16. Howley ET, Bassett DR, Welch HG. Criteria for maximal oxygen uptake: Review and commentary. Med. Sci. Sports Exerc. 1995; 27: 1292-1301.

17. Ellestad MS. Stress Testing Principles and Practice (3rd Edition). Philadelphia, F.A. Davis. 1986.

18. Bar-Or O. A new anaerobic capacity test: characteristics and applications. Proceedings of the 21st World Congress in Sports medicine. Sep 7-12, 1978, Brasilia.

19. Bar-Or O. The Wingate anaerobic test: an update on methodology, reliability and validity. Sports Medicine. 1987; 4: 381-394.

20. Wilmore JH, Vodak PA, Parr RB, Girandola RN, Billing JE. Further simplification of a method for determination of residual lung volume. Med. Sci. Sports Exerc. 12; 1980: 216-218.

21. Jackson AS, Pollock ML. Generalized equations for predicting body density of men. Br. J. Clin. Nutr. 1978; 40: 497-504.

22. Jackson AS, Pollock ML, Ward A. Generalized equations for prediction body density of women. Med. Sci. Sports Exerc. 1980; 12: 175-182.

IN VITRO EVALUATION OF LUMPED PARAMETER ARTERIAL MODELS OF THE CARDIOVASCULAR SYSTEM

Jeremy D. Schaub
Graduate Research Assistant
Microelectronics Research Center
Department of Electrical Engineering

University of Texas at Austin
Austin, TX 78722

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

September 1996

IN VITRO EVALUATION OF LUMPED PARAMETER ARTERIAL MODELS OF THE CARDIOVASCULAR SYSTEM

Jeremy D. Schaub
Graduate Research Assistant
Microelectronics Research Center
Department of Electrical Engineering
University of Texas at Austin

Abstract

Five electrical analog models of the arterial system were tested for accuracy in predicting systemic arterial compliance (SAC) and total peripheral resistance (TPR). Aortic pressure and flow waveforms were generated using a cardiovascular dynamics simulation modeling (CDSM) system at heart rates and stroke volumes typically seen in rhesus monkeys while the compliance and resistance of the system was simultaneously measured. All five models predicted TPR values 24-29% higher than the measured resistance value due to a non-linear screw clamp resistor used in the experimental setup. The four-element Noordergraaf model produced the most accurate estimates (10% percent error) for SAC, while the inductance and Windkessel models produced errors of 16% and 54%, respectively. The Westkessel and Schroeder models encountered problems due to the high heart rates seen experimentally and reproduced in this study.

IN VITRO EVALUATION OF LUMPED PARAMETER ARTERIAL MODELS OF THE CARDIOVASCULAR SYSTEM

Jeremy D. Schaub

Introduction

Many lumped parameter electrical analog models of the arterial system [1-5] have been used to provide beat-to-beat estimates of systemic arterial compliance (SAC) and total peripheral resistance (TPR) given aortic pressure and flow as inputs to the system. The accuracy of these models has been evaluated using a variety of techniques, including comparing aortic input impedance (magnitude and phase) of the model versus *in vivo* data obtained experimentally [6]. Another technique involves injecting aortic flow into the model and comparing the resulting pressure waveform to an experimentally obtained aortic pressure. A problem with these techniques is that the predicted values for SAC and TPR cannot be compared to experimental values, since direct measurement of these *in vivo* parameters can be difficult. An *in vitro* pulsatile circulatory system that simulates aortic pressure and flow characteristics of physiologic data provides a solution to this problem. The resistance and compliance of the system can be characterized completely, and these values can be directly compared to the SAC and TPR predicted by the arterial models. The purpose of this study was to evaluate the accuracy of five arterial models in predicting SAC and TPR using a cardiovascular pulse duplicator simulation system.

Methodology

Arterial Models

The arterial models tested in this study are shown in Figure 1. The two-element Windkessel model (Figure 1a) consists of a capacitor (C) and resistor (R) in parallel, representing systemic arterial compliance (SAC) and total peripheral resistance (TPR) for the entire systemic vasculature [1]. The 3-element Westkessel model (Figure 1b) includes a resistance (Z_o) term, representing proximal arterial (aortic) impedance, in series with the traditional Windkessel model [2]. The 3-element inductance model (Figure 1c) replaces the series resistance with an inductor (L) to represent inertia of blood through the aorta. The parameters of the 4-element Noordergraaf model (Figure 1d) represent resistance (Z_o) and inertia (L) of blood through the aorta, compliance (C) of the elastic wall, and viscoelasticity or leakage flow through branches (R coupled with C) [3]. The 4-element Schroeder model (Figure

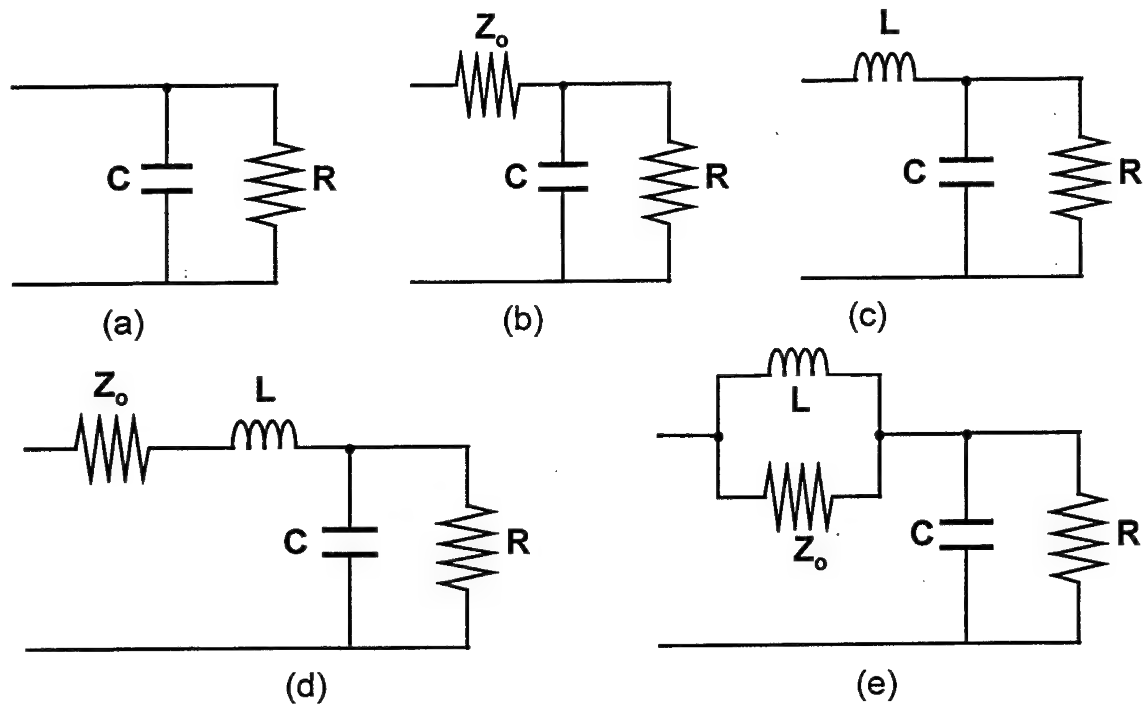


Figure 1: Arterial models tested in the study. (a) 2-element Windkessel [1], (b) 3-element Westkessel [2] , (c) 3-element inductance model ,(d) 4-element Noordergraaf [3], and (e) 4-element Schroeder [4].

1e) is a modified version of the 4-element Noordergraaf model, in which the inductor (L) was placed in parallel with the characteristic impedance (Z_o) [4].

Parameter estimates for the Windkessel model were obtained using the method proposed by Self et al. [1]. The other four models estimated SAC and TPR using a modified version of Newton's method to obtain convergence between estimated and *in vivo* impedance coefficients [5].

Setup

Arterial models were tested using simulated aortic pressure and flow profiles generated by a cardiovascular dynamics simulation modeling (CDSM) system [7]. The system was optimized to simulate the aortic impedance characteristics (magnitude and phase) of data obtained experimentally in a rhesus monkey. Compliance and resistance of the system were accomplished with a single compliant tube (7.5" length, 0.62" diameter, 0.045" wall thickness) and screw clamp resistor, as shown in Figure 2. In-line pressure/flow modules [8] were placed both upstream and downstream of the compliant tube for simultaneous measurement of pressure and flow. Data were

collected at heart rates of 130, 150 and 170 beats per minute and at stroke volumes of 4, 6 and 8 cc, which correspond to the experimentally obtained data. [9].

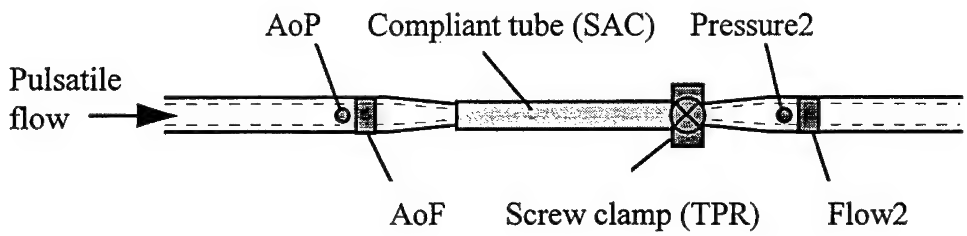


Figure 2: Compliance and resistance model test section used simulate experimental data.

Data Analysis

The simulated values for resistance were calculated on a beat-to-beat basis using Equation 1,

$$\text{Resistance} = \frac{\text{mean}(\text{AoP}) - \text{mean}(\text{Pressure2})}{\text{mean}(\text{AoF})}, \tag{1}$$

Compliance was defined as,

$$\text{Compliance} = \frac{\Delta\text{Volume}}{\Delta\text{Pressure}}. \tag{2}$$

The change in volume of the compliant tube was found by integrating the two flow waveforms,

$$\Delta\text{Volume} = \int(\text{AoF} - \text{Flow2})dt, \tag{3}$$

which was evaluated on a point-to-point basis to get the volume as a function of time. The change in pressure in Equation 2 refers to the pressure inside the compliant tube. However, only aortic pressure was measured during normal data collection. A preliminary test was conducted to characterize the relationship between aortic and tube pressure. A Millar® pressure catheter was inserted into the tube and a valve was shut downstream of the tube to prevent any fluid from escaping. Fluid was then injected into the tube while pressure and flow profiles were recorded. Typical waveforms from this test are shown in Figure 3a, and the pressure-volume relationship is shown in Figure 3b.

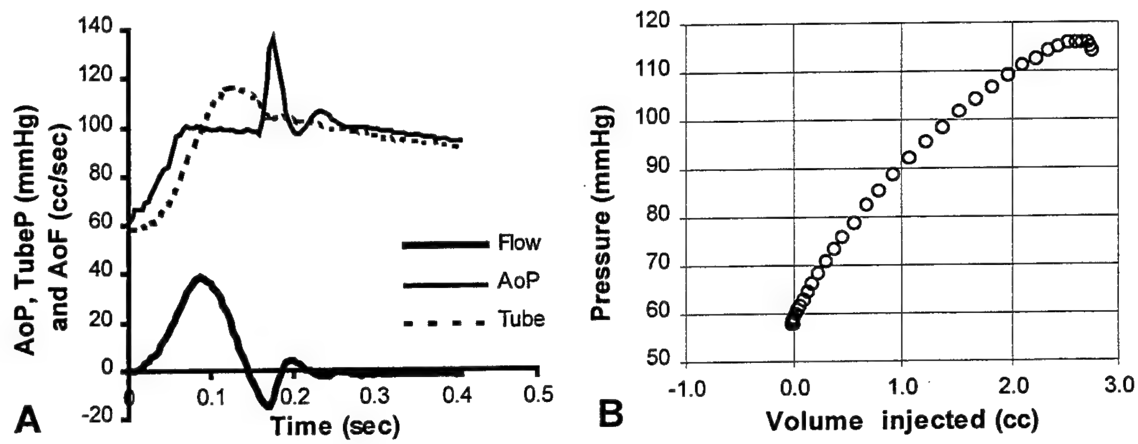


Figure 3: Preliminary test waveforms: (A) simulated pressure and flow profiles, and (B) corresponding pressure-volume relationship.

The phase delay and difference in peak magnitude between aortic and tube pressure in Figure 3a indicate that the inertance of the fluid was significant. The relationship between aortic and tube pressure was defined as,

$$\text{TubeP} = \text{AoP} - L\frac{d(\text{AoF})}{dt}, \tag{4}$$

where L is the inertance of the fluid. To obtain a value for this inertance, Equation 4 was solved on a point-by-point basis and the mean value for L was calculated.

The ideal pressure-volume relationship (Figure 3b) should be a straight line, as indicated by Equation 2. To correct for the non-linearity seen in the preliminary test, the compliant tube was modeled as a first order system with both resistance (R') and compliance (SAC). The equivalent electrical circuit is shown in Figure 4.

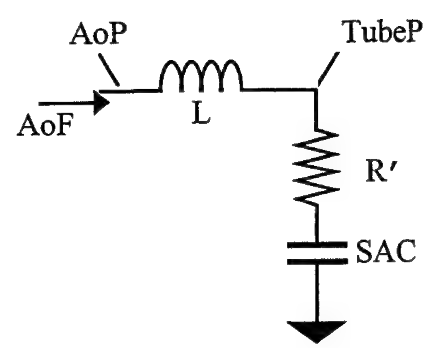


Figure 4: Electrical circuit model of preliminary test setup.

From this model, the tube pressure was calculated using Equation 5,

$$\text{TubeP} = \text{AoF} \times R' - \frac{1}{\text{SAC}} \times \text{Vol} + P_i, \tag{5}$$

where P_i was the pressure at the beginning of injection. During data analysis, the tube pressure was calculated using both Equation 4 and

Equation 5. The R′ and SAC values were adjusted in Equation 5 until the norm-squared error between the two pressures was minimized. This 'true' compliance value was then compared to the compliance predictions of the arterial models.

Results

All of the models produced predictions for SAC and TPR except the Schroeder model, which did not converge for heart rates of 150 and 170 beats per minute (subsequently, these results were omitted). The SAC and TPR estimates of the Schroeder model that did converge were slightly less accurate than those of the Noordergraaf model, but are not presented.

HR (bpm)	Windkessel	Inductance	Westkessel	Noordergraaf
130	1.4	1.6	1.4	1.5
150	0.5	0.7	0.6	0.6
170	0.4	0.6	1.3	0.7
ALL	0.8	1.0	1.1	0.9

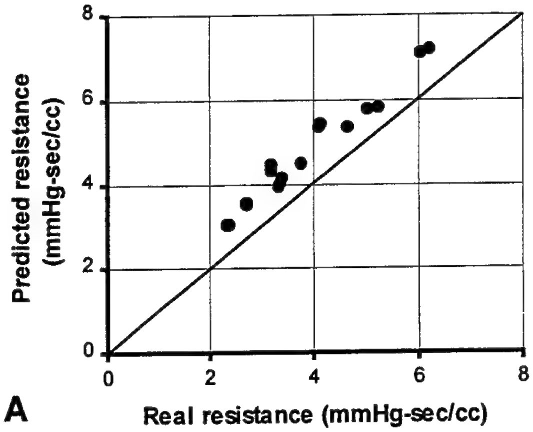
Table 1: Mean square error between true and predicted TPR values of the four arterial models.

The mean square error of the TPR estimates are given in Table 1. The mean square error of 0.8 $mmHg^2$-sec^2/cc^2 between true and predicted values was smallest for the Windkessel model. This corresponds to a percent error of 24%. Figure 5a shows the true and estimated values of resistance for the Windkessel model.

HR (bpm)	Windkessel	Inductance	Westkessel	Noordergraaf
130	6.1	0.4	80	0.0
150	12	1.1	500	0.5
170	14	1.5	27000	1.1
ALL	11	1.0	8100	0.5

Table 2: Mean square error * 10^4 between true and predicted SAC values

The mean square error of the SAC estimates are given in Table 2. The mean square error of $0.5*10^{-4}$ $cc^2/mmHg^2$, corresponding to a percent error of 10%, was smallest for the Noordergraaf model. The SAC predictions of this model are plotted versus the true values in Figure 5b.



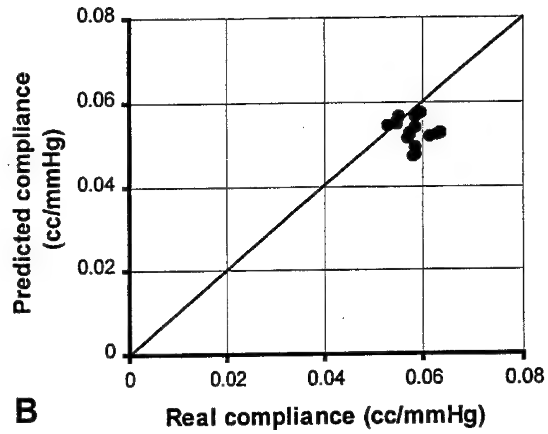


Figure 5: (A) Windkessel predictions for TPR and (B) Noordergraaf predictions for SAC

Discussion

The estimates of TPR were similar between all five arterial models tested in this study. However, each model exhibited an offset shift, as seen in Figure 5a. This shift was most likely caused by the screw clamp, which increases in resistance non-linearly as the pressure increases. Thus, the resistance may have changed within a beat due to the pulse pressure ($\Delta P \approx 70$ mmHg) produced by the CDSM system. Since each arterial model assumed that TPR was constant over each beat, the predicted values were higher than those calculated using Equation 1. Since the pulse pressure increased as the heart rate declined, the mean square error was largest at 130 beats per minute for each model. To alleviate these problems, a linear resistor could be inserted into the experimental setup or Equation 1 could be modified to account for the distortion.

The true values for compliance remained fairly constant at each heart rate and stroke volume setting, as seen in Figure 5b. This indicates that the first order approximation used to calculate the true value for SAC was effective. However, the compliance predictions (Table 2) show considerably more spread than in the TPR estimates, as the mean square error of the Westkessel predictions was over four orders of magnitude higher than the Noordergraaf estimates of SAC. In addition, the compliance predictions worsened as the heart rate increased, which may have been a result of simulating high frequency data. The impedance of a hydrodynamic system is generally more inductive at higher frequencies, therefore it was not surprising that the inductance and Noordergraaf models produced the most accurate compliance estimates. The lack of an inertial component in the Westkessel and Windkessel models most likely attributed to their poor compliance estimates. The higher frequency data may

further explain the convergence problems encountered using the Schroeder model, where 93-99% convergence was achieved using lower frequency data obtained experimentally in baboons [5].

Conclusions

The estimates of SAC and TPR for five arterial models was tested and evaluated using a cardiovascular pulse duplicator system that simulated physiologically equivalent arterial parameters of experimentally obtained data in a rhesus monkey. No discernible differences in the TPR predictions between the models was detected, however, each of the models overestimated the true resistance by approximately 25%. This overestimation was probably due to the non-linear properties of the screw clamp used to simulate the true resistance. The four-element Noordergraaf model produced the most accurate estimates for SAC, while the inductance and Windkessel models produced errors of 16% and 54%, respectively. The Westkessel and Schroeder models had problems estimating compliance most likely due the higher hearts seen experimentally in rhesus monkeys and simulated in this study.

References

[1] Self, A, DL Ewert, RD Swope, R Crisman, and RD Latham. 'A new method for beat-to-beat determination of peripheral resistance and systemic arterial compliance under non-steady-state conditions during +Gz centrifugation'. Aviat. Space Environ. Med., 65:396-403, 1994.

[2] Westerhoff, N, G Elzinga, and P Sipkema. 'An artificial arterial system for pumping hearts'. J Appl. Physiol., 31:776-781, 1971.

[3] Milnor, W. Hemodynamics, 2nd Edition. Baltimore: MD, Williams and Wilkins, 1989.

[4] Schroeder, MJ. Optimal Ventricular-Arterial Coupling with Respect to Arterial Elastance. MS Thesis, North Dakota State University, Fargo, N.D., 1994.

[5] Essler, SJ. Estimation of Arterial Vascular Parameters for Transient and Steady Beats. MS Thesis, North Dakota State University, Fargo, N.D., 1995.

[6] Koenig SC, DL Ewert, JD Schaub, RD Swope, V Convertino, JW Valvano. 'Evaluation of electrical analog models of the arterial system.' in preparation.

[7] Schaub, JD, SC Koenig, CD Offerdahl, DL Ewert, RD Swope. 'Cardiovascular dynamics simulation modeling system.' Med. & Biol. Eng. & Comput., in review.

[8] Koenig, SC, JD Schaub, DL Ewert, RD Swope. 'In-line pressure-flow module for cardiovascular dynamics in vitro modeling and biosensor validation.' Med. & Biol. Eng. & Comput, in review.

[9] Koenig, SC. Modeling mechanisms of alterations in central venous pressure during exposure to 10° head down tilt in a chronically-instrumented non-human primate. PhD Dissertation, University of Texas, Austin, TX. 1996.

JAVA-BASED APPLICATION OF THE MODEL-VIEW-CONTROLLER FRAMEWORK IN DEVELOPING INTERFACES TO INTERACTIVE SIMULATIONS

Nicole L. Schneider
Graduate Teaching Assistant
Department of Biomedical and Human Factors Engineering

207 Russ Engineering Center
College of Engineering and Computer Science
Wright State University
Dayton, OH 45435
Tel: +1 513 873 5071
Fax: +1 513 873 5009
Email: snarayan@cs.wright.edu

Final Report for:
Summer Faculty Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Armstrong Laboratory

September 1996

JAVA-BASED APPLICATION OF THE MODEL-VIEW-CONTROLLER FRAMEWORK IN DEVELOPING INTERFACES TO INTERACTIVE SIMULATIONS

S. Narayanan and Nicole L. Schneider
Assistant Professor
Department of Biomedical and Human Factors Engineering
Wright State University
Graduate Teaching Assistant
Department of Biomedical and Human Factors Engineering
Wright State University

Abstract

Interfaces to simulations serve to portray the dynamic behavior of the modeled system. In visual interactive simulations, user interfaces allow an analyst to also interact actively with the executing simulation. Traditionally, the software to display the simulation model and to facilitate user interaction are embedded in the simulation model. Such an integration makes it difficult to maintain large simulation programs and pose limitations in the development of multiple interfaces to a simulation model. This article presents a Java-Based Architecture for Developing Interactive Simulations (JADIS). JADIS applies the Model-View-Controller paradigm to the development of interactive simulations. In JADIS, the simulation model and multiple interfaces to them are separate processes that execute concurrently on distributed machines. JADIS integrates concepts from object-oriented programming, concurrent, distributed processing, and graphical user interface design in developing visual interactive simulations. This article describes the JADIS architecture and presents application of JADIS to the airbase logistics modeling domain.

JAVA-BASED APPLICATION OF THE MODEL-VIEW-CONTROLLER FRAMEWORK IN DEVELOPING INTERFACES TO INTERACTIVE SIMULATIONS

S. Narayanan and Nicole L. Schneider

Introduction

Discrete-event simulations offer flexible and powerful means of observing and analyzing complex, dynamic systems. The fundamental objective in simulation is to use the software abstractions provided by the language to represent the behavior of system entities over time. In traditional discrete-event simulations, the analyst is primarily a passive observer of the simulation program execution. With the increase in computing power and graphical user interfaces, there is an increasing interest in the area of visual interactive simulations (VIS)[Bell & O'Keefe, 1987; Bell, 1991; Hurrion, 1980; Lyu & Gunasekaran, 1993; McGregor & Randhawa, 1994].

In VIS, interfaces serve to not only display the state of the simulated system, but also to allow an analyst to interact with the executing simulation. The simulation executes in real-time or scaled time. The analyst can modify the parameters of the simulation, alter the dynamics of the simulated system, and can pause/restart the simulation. The VIS approach offers several potential advantages. First, it allows the user to make complex decisions. For example, Hurrion and Secker (1978) found that the rules used by human schedulers in job shop scheduling were difficult to encapsulate in simulation models. VIS offered a viable alternative. Second, VIS are useful in studying the effectiveness of real-time, human decision making in complex systems. Dunkler, et al. (1988), for example, used an interactive simulation of a flexible manufacturing system and compared the effectiveness of various automatic scheduling strategies with that of human scheduling in expediting parts through the system. Third, the display of the simulated system in VIS can be visually appealing and can increase effective communication between a manager and the simulation analyst in model development (Bell, 1991; Bishop & Balci, 1990). Fourth, the dynamic visual representation in VIS can highlight logical inconsistencies in the model and can therefore be effective in model verification and validation. Finally, since the user of VIS actively participates in the execution of the simulation, there is potential for increased user confidence in applying the results of the simulation (Kirkpatrick & Bell, 1989).

There are several potential problems with the VIS approach (Bell & O'Keefe, 1987; Bishop & Balci, 1990; Paul, 1989). First, due to human interaction at various times during the execution of the simulation, the simulation experiments are hard to duplicate and are not amenable to traditional simulation output statistical analysis. Second, a user interacting with the simulation may observe a snapshot of the system and may prematurely conclude that the system will always exhibit the observed characteristics without the benefit of detailed analysis.

Third, in the VIS approach, design of the dynamic display tends to be an integral part of the simulation model development making the traditional simulation life cycle inadequate to describe the VIS approach (Hurrion & Secker, 1978).

Despite the problems outlined above, when applied appropriately, interactive simulations are useful in the analysis of complex, dynamic systems. They are necessary to analyze human interaction with complex systems and can be effective in enhancing user understanding of large, semi-structured problems through interaction with the simulation.

The major challenges in developing interactive simulations are problems associated with computer hardware and software (Bell & O'Keefe, 1987). Bell (1991) highlights the historic struggle of the early VIS development effort with advances in computer hardware. Early VIS systems including See-Why were developed for large main frames. Currently, personal computers and workstations have become the standard for systems development. Most VIS packages currently available are still hardware dependent and suffer from problems of portability.

Several early interactive simulation packages were developed in FORTRAN (e.g., FORSSIGHT). Developmental interest has moved towards C and recently towards object-oriented languages (e.g., ProfiSEE in Smalltalk-80 [Vaessen, 1989]). While object-oriented programming offers many advantages for simulation modeling in terms of modularity, software reuse, and natural mapping with real world entities (Narayanan et al., 1996), their application to developing interactive simulations has been only explored in a limited way (Bell, 1991). The software to display the simulation model and to facilitate user interaction are embedded in the simulation model. Such an integration makes it difficult to maintain large simulation programs and pose limitations in the development of multiple interfaces to a simulation model. Although it is acknowledged that the interface configuration and interaction specification are concurrent with the model specification, effective means to facilitate concurrent software development are lacking. As a result of the tight coupling of the simulation model with the interface, it is often difficult for concurrent development of the two phases.

This article presents a Java-Based Architecture for Developing Interactive Simulations (JADIS). JADIS applies the Model-View-Controller paradigm from Smalltalk to the development of interactive simulations (Goldberg, 1990). In JADIS, the simulation model and multiple interfaces to them are separate processes that execute concurrently on distributed machines. JADIS integrates object-oriented programming, concurrent, distributed processing, and graphical user interface design in developing visual interactive simulations.

The remainder of the article is organized as follows. First, we present background to visual interactive simulations and the model-view-controller framework. We then discuss the motivations for applying the MVC framework to developing interfaces to interactive simulations and present the JADIS architecture. We will then describe the application of JADIS to airbase logistics. We will discuss our approach in the context of existing efforts in the literature and conclude with a summary of contributions of this study and suggest recommendations for future work.

Background

Table 1 presents a comparison of interactive simulations with traditional discrete-event simulations and animated simulations along seven dimensions: nature of suitable problems, simulation development life cycle, time transition of simulation clock, nature of user interaction, role of the graphical interface, types of output analysis, and example of software packages for each category. Interactive simulations are well suited for large, semi-structured problems in which human interaction is an important consideration. Interactive simulation development is different from the traditional simulation life cycle as the specification of interaction and animation is concurrent with model specification. The simulation clock is updated either on a real-time basis or on a scaled time. The graphical interface in interactive simulations depict dynamic system states, highlight performance measures, and contain interface objects that accommodate command line inputs and other user interaction. Output analysis in interactive simulations primarily involves transient systems analysis.

Table 1. Comparison of Interactive Simulations with Traditional Discrete-Event Simulations and Animated Simulations.

Topic/Issue	**Traditional discrete-event simulations**	**Animated discrete-event simulations**	**Interactive simulations**
Nature of suitable problems	• Well-structured problems • Small to medium scale systems • Human interaction not a critical consideration	• Well-structured to semi-structured problems • Medium to large scale systems • Human interaction not a critical consideration or can be captured completely	• Semi-structured to unstructured problems • Small to medium scale systems • Human interaction is a critical consideration
Simulation development life cycle	• Traditional	• Traditional with animation specification following simulation model configuration	• Interface and interaction specification concurrent with simulation model specification
Time transition of simulation clock	• Discrete-event to another	• Discrete-event to another with scaled time for animation	• Discrete-event to another • Scaled time • Real time

User interaction	• None	• None or able to alter simulation experimental parameters	• Able to alter simulation experimental parameters • Able to alter simulation system dynamics
Graphical interface	• None	• Displays dynamic system states • Displays performance measures	• Displays dynamic system states • Displays performance measures • Interface objects accommodate command line input and other user interaction
Output Analysis	• Conventional steady state or terminating simulation analysis	• Conventional steady state or terminating simulation analysis	• Transient systems analysis • Human factors engineering analysis methods
Software	• Simulation languages (e.g., GPSS V) • Programming language (e.g., FORTRAN)	• Simulation packages with animation capability (e.g., SLAM/TESS, SIMAN/CINEMA) • Programming language (e.g., FORTRAN, C)	• Simulation packages (e.g., VISION, SEE-WHY, WITNESS) • Programming language (e.g., C, C++, Smalltalk)

The major challenge in developing visual interactive simulations is associated with computer hardware and software problems (Bell, 1991). There is need for computational architectures that can enable the development of interactive simulations hardware independent. Also, object-oriented languages offer tremendous promise and are obvious vehicles for VIS development (Bell and O'Keefe, 1997). Several packages such as Audition and ProfiSEE (Vaessen, 1989) have been developed using an object-oriented language. Existing packages, however, exploit the power of object-based programming and concepts such as model-view-controller framework in only a limited manner. Most architectures are also hardware dependent.

This article discusses a Java-based Architecture for Developing Interactive Simulations (JADIS). JADIS overcomes the hardware and software limitations of traditional VIS architectures outlined above. JADIS is developed using the Java programming language (Lemay & Perkins, 1996). The Java source code is compiled into byte code that can be read by an interpreter available on multiple platforms including personal computers, Macintoshes, and UNIX workstations. The software can be developed on any platform that contains the Java Development Kit (JDK). JDK is available on most operating systems. The byte code can then be moved to another platform and can be run successfully without altering a single line of code. Figure 1 illustrates the Java virtual machine environment.

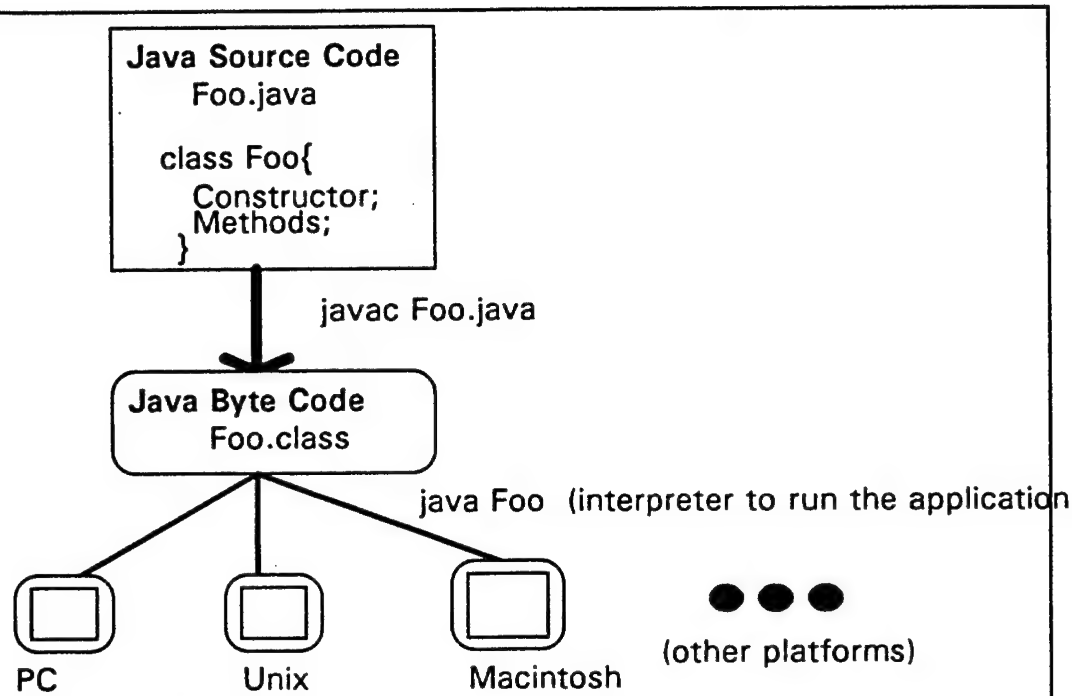


Figure 1. Illustration of the Java virtual machine environment.

Java is an object-oriented programming language whose syntax is similar to C++. Java supports encapsulation, inheritance, and polymorphism. It, however, does not have explicit pointers, doesn't support multiple inheritance, and doesn't feature operator overloading. Java's popularity is initially because of applets, which are written in Java and can be embedded on home pages on the world wide web. Applets add animation and interaction to web pages and can be viewed using a browser such as Netscape 2.0 or higher. Java can also be used as a regular programming language where application can stand alone without being embedded as applets. The Java language comes with various packages (similar to libraries) for general data structures, applets, file input/output, and also for graphical user interfaces. Java is multithreaded and hence particularly suitable for distributed computing as it easily copes with TCP/IP protocols. Java can thus be used for both creating simulations as well as for creating interfaces to simulations. The Java language also features a utility called javadoc which enables automatic hypertext generation of software documentation. Users can add specialized comments in the source code which can be easily processed by javadoc in generation of source code documentation.

JADIS applies the Model-View-Controller (MVC) paradigm from Smalltalk to the development of interactive simulations. We first describe the MVC paradigm before describing JADIS.

MVC is a paradigm for developing graphical user interfaces in a modular manner (Gobbetti and Turner, 1991, Goldberg, 1990; Krasner and Pope, 1988). In MVC, any interactive program is conceptually divided into three areas: (1) the Model, which contains representation of the application domain, (2) the View, which contains specification of the display, and (3) the Controller, which contains specification of user interactions with the underlying model. In the context of interactive simulations, the model refers to the simulation model, the view corresponds to how the dynamics of the simulation are displayed to the user, and controller refers to the processing of user commands input to the simulation model through the display.

The MVC framework provides several potential advantages in developing interfaces to interactive simulations. First, due to the separation of the model from the view, the simulation model development can take place concurrently with the specification of the interface. The simulation developer can focus on the model development and leave the responsibility of the interface design to a graphical user interface (GUI) designer. Second, multiple views to the same simulation can be developed. The end user can then plug different displays or pieces of code into the simulation. The simulation model can be arranged to suit the needs of individual modelers without requiring programmers to constantly create entirely new code. Thus, the productivity of software development is enhanced. The reuse of existing designs and refinements potentially also leads to stable applications with a consistent style.

MVC is an improvement over previous approaches to developing interactive simulations. Simulation modelers need no longer be experts at implementing simulation models as well as be able to design and implement graphical displays. GUI experts can create display modules which can be stored in graphical libraries for the simulation analysts to use in customizing the simulation view. With MVC, many users can access multiple, simultaneous views of the same simulation model.

The JADIS Architecture

Figure 2 presents a schematic of the JADIS architecture. There are three primary modules in JADIS: (1) simulation module comprising the simulation infrastructure including clock, random number generators, various statistical distributions, and event calendar; and domain-specific classes consisting of general queuing utilities and system-related classes (for example in air base logistics simulation classes include air base, hangar, and air craft); (2) interface module consisting of general display classes (e.g., list boxes) and domain-specific display elements (e.g., air craft), and (3) inter-process communication module consisting of sockets and the ability of the simulation to broadcast state changes to the views and the ability of the interface process to send commands to the simulation and receive messages from the simulation. The simulation process instantiated from the simulation module and the interface process instantiated from the interface module can run concurrently on the same machine or on different

platforms. Users can input commands to the simulation through the command line or through the interface. Messages travel between the simulation and the interface processes in the form of text strings.

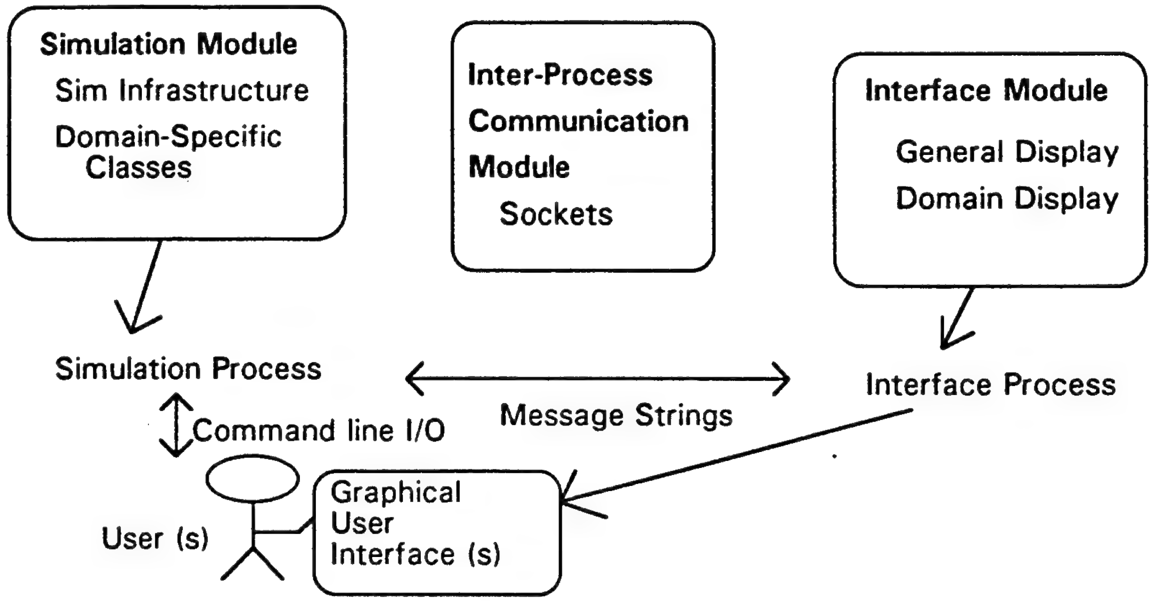


Figure 2. Schematic of the JADIS Architecture.

The JADIS architecture alters the simulation model development process. In using JADIS, the analyst defines the elements of the simulation model. The analyst then defines the interface and specifies the communication messages. The simulation model and views are thus concurrently specified. The analyst creates instances or subclasses as appropriate and interconnects the simulation and interface processes. The software is then tested, model is verified, and hypertext source code documentation is automatically generated through the `javadoc` utility in Java. The JADIS architecture has been applied in the interactive simulation of airbase logistics. The application is discussed in the next section.

Application to Airbase Logistics

The domain of airbase logistics is large and complex. It involves logistics processes that support aircraft sortie generation at operational airbases. Airbase logistics involves aircraft maintenance, parts supply, and munitions loading (Popken, 1992). Models of logistics processes are useful in analysis for aircraft acquisition planning, maintenance manpower allocation, and theater-level supply redistribution. Popken (1992) discusses that a synergistic combination of object-oriented programming, databases, and graphical user interfaces provide

significant enhancements to simulation modeling capabilities. An aircraft maintenance problem provided an excellent testbed for demonstrating the JADIS architecture.

This section is organized as follows. First, we outline the features of the system and the assumptions made. We then describe the principles in designing the simulation model and the interface. We discuss the various classes and highlight the salient class hierarchies, and detail the capabilities of the system. We also describe the lessons learned during the model development process.

The model used is designed to capture maintenance operations at an airbase. There are aircraft of different kinds with varying configurations and capabilities. An aircraft is comprised of several subsystems. Sorties are generated by different methods (e.g., random generation). Each sortie specifies the number of aircraft required, the type of each aircraft, and the details of the mission. While the aircraft is in operation one or more of its subsystems may fail. When a subsystem of an aircraft fails, it is sent to the maintenance facility for repairs. The maintenance facility includes a hangar, different types of *test* equipment, spare parts, and personnel. Various performance measures in this system include maintenance cost, sorties completed, sorties aborted, hangar utilization, and personnel utilization. We developed classes to represent the entities in this system and to specify their interaction (Carrico & Clark, 1995; Carrico et al., 1995).

The specific implementation made some simplifying assumptions. First, the model was at the airbase level and not at the theater level. Second, the maintenance resources were always available. Third, subsystems featured a single failure. Each subsystem failure identified the unique maintenance actions required. Sorties are generated randomly, in a pre-set mode, or in the fly-when-ready mode. The simulation duration was two weeks, with sortie generation occurring 16 hours each day, seven days a week.

Several design principles were applied both in the simulation model development and in the interface design. First, we exploited the capability of natural mapping and modularity features of object-oriented programming. Through object-oriented programming it is possible to develop software abstractions that have a direct correspondence with real world objects (Narayanan et al., 1996). Objects can also be reused through inheritance. Second, in the simulation model, physical objects (e.g., aircraft, subsystem, and hangar) were distinguished from decision making objects (e.g., scheduler, resource manager) and information storage objects (e.g., resource statistics). The advantage of making such a distinction is to allow different decision making strategies to be evaluated using the simulation model, where only the decision making entities need to be changed. Third, the interactions between objects were limited thereby enhancing the plug-in capability of the architecture.

On the interface development process, information was presented hierarchically. The analyst can get to see the overall dynamics of the simulated system and if desired could then look at details of the individual components. Second, the interface accommodates users in multiple modes where the users interest may be in systems analysis or simply in visualizing the maintenance processes. Third, interfaces feature a standard look and feel as that of Motif. It also features balloon help as in Microsoft Windows. Finally, the components of the architecture for interface development can be easily assembled for rapidly prototyping graphical user interfaces for a class of similar problems.

The application developed consists of two distinct processes: (1) simulation and (2) interface. These processes can either run on the same machine or different machines. All simulation classes are descendants of `SimBase`. Figure 3 highlights salient top-level classes in the simulation model. Two major subclasses of SimBase are `ActiveSimulationObject` and `InformationStore`. `DecisionMaker` and `Physical` are two subclasses of `ActiveSimulationObject`. Subclasses of `InformationStore` include `ResourceStatistics` and `MaintenanceInfo`. Similarly, `Equipment`, `Personnel`, and `Hangar` are all subclasses of `Resource`. DecisionMakers include `ResourceManager` and `Coordinator`. All of these classes have a natural mapping to real world entities.

On the interface process, major classes include **`Interface`** which initiates threads for receiving input from simulation and for setting up the animation, class `Animation` (inherited from Java's `Frame` class) which sets up displays by instantiating the domain specific display objects such as hangar, aircraft, and runway, class `CommandEntry` which facilitates user interaction in a command window, class `Dynamics` which displays visualization of the processes involved in airbase maintenance, class `EventList` which displays a log of events as they occur on the simulation side, classes `GraphSS` and `Graphs` to display graphs of various performance measures, and classes for menus, push button, etc. overloaded from Java's abstract windowing toolkit. **`Animation`** has a **`processEvent`** method which in turn invokes `processEvent` of displayed objects. The knowledge of how displayed objects update to the events in simulation are encapsulated in the class representation of the displayed objects.

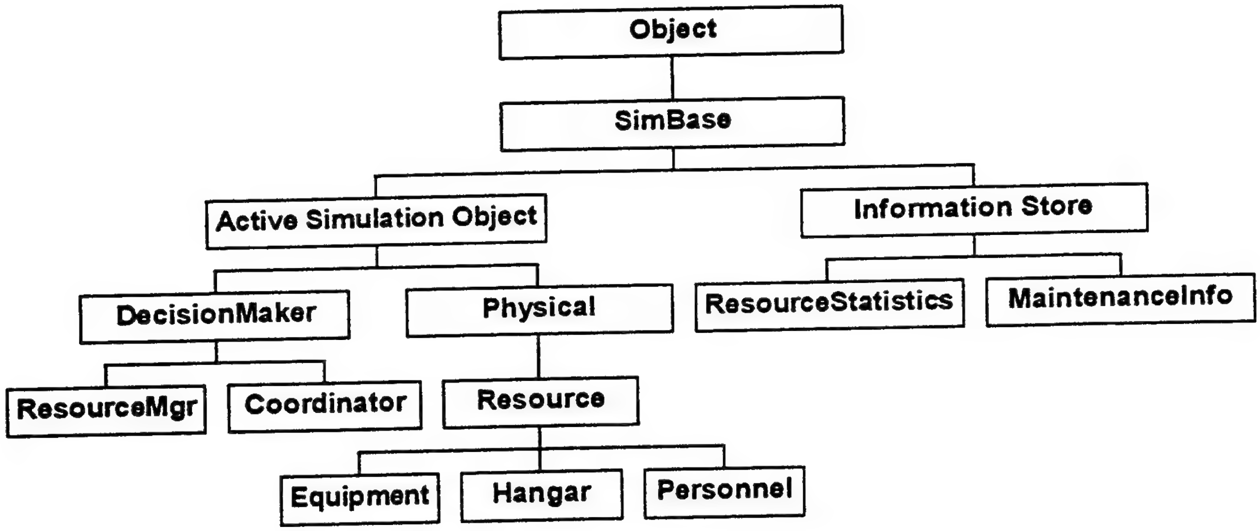


Figure 3. Illustrative hierarchy of simulation classes to model aircraft maintenance.

There were a total of 49 Java classes implemented for the simulation model of this application. There were 15 additional classes used for the interface. The size of classes ranged from 20 to over 800 lines totalling approximately 21,000 lines of code. Screen output of the main interface is included in the appendix. The javadoc capability of Java was used in automatically generating the source code documentation with hypertext links. Javadoc turns comments to code into hypertext markup language code with links to the parent class and overridden methods. The URL for the source code documentation of the application is `http://isis.cs.wright.edu:1947/fx99html/`. A sample javadoc generated file is included in the appendix for illustrative purposes.

Discussion

This report presented JADIS, a Java-based architecture for developing interactive simulations. The JADIS architecture integrated concepts from object-oriented programming, distributed computing, and graphical user interfaces to interactive simulations. Since the architecture is implemented in Java, it is hardware independent. The JADIS architecture is an instantiation of the Model-View-Controller framework in interactive simulations.

JADIS, however, goes beyond the traditional implementation of the MVC framework. While the MVC framework provides a powerful metaphor for developing interactive simulations, practical implementations often lead to complicated, unwieldy class inheritance structures (Krasner & Pope, 1988; Shan, 1990). In JADIS, the model and views are completely separated. The inheritance structure maps well to the real world objects. The semantics of how displayed objects are updated is encapsulated well within the class representation. Also, in the traditional

MVC, a model broadcasts that its status has changed, all views and controllers tied to that model are required to query it to discover exactly what the change is before they can update themselves. JADIS reduces redundancy in message passing by having the simulation model broadcast that its state has changed and the details of the state change to the views. Finally, in contrast to the polling protocol applied in the traditional MVC, JADIS applies an process-based, event-driven protocol which maps well with simulations where behavior is represented as events occurring at different time units.

The architecture also facilitates the instantiation of interactive simulations. This capability goes well beyond animating discrete-event simulations such as seen in MicroSaint or CINEMA for SIMAN. In JADIS simulations, users can alter the parameters of the simulation and also modify the system dynamics. For example, users in the airbase logistics simulation can alter the parameters of the maintenance resources at run time and also alter the sortie generation discipline. Real-time human decision making can therefore be readily studied using JADIS simulations. The ability to run the simulation and interface on multiple machines concurrently is also a powerful capability in harnessing the power of distributed computing.

The architecture currently has a few limitations. First, it has not yet been tested for handling simultaneous user input from multiple views of the same simulation model. Second, the simulation and the interface are implemented as Java applications rather than Java applets. Therefore, they can not directly be viewed using an internet browser. Third, when the number of messages between the simulation and the interface becomes very high, it slows down the machine. This limitation can be overcome by incorporating capabilities to filter needless messages appropriately. The airbase modeling application has some limitations as well. First, the model was developed at the airbase level and not at the theater level. Second, human interaction was limited to altering clock speeds, viewing performance measures, altering maintenance data files, and scheduling disciplines of sorties dynamically. Third, the maintenance behaviors were also simplified to feature single failure subsystem and also having adequate spares and other maintenance resources.

Future research extensions include incorporating the capability to run the simulation and view it on internet, extending the scope of the application to include modeling at the theater level, incorporating additional human interaction, enhancing the visualization capabilities in the system, and empirically evaluating the efficacy of interfaces tailored to users. The ultimate goal is to have a high-fidelity computational representation of airbase logistics so as to support logistical decision making through computer-based tools. Integrating interactive optimization capabilities to the descriptive simulation modeling architecture is another promising avenue of research.

Conclusions

Advances in software offer unique opportunities in enhancing simulation modeling capabilities. Interactive simulation is a useful methodology for systems analysis of large, complex systems. Traditionally, the development of interactive systems have been plagued with software and hardware problems. We have applied Java programming language and integrated concepts from object-oriented programming, model-view-controller, distributed computing to develop a JADIS, a Java-Based Architecture for Developing Interactive Simulations.

The JADIS architecture can run on any platform which supports the Java Development Kit. Such systems include PCs running Windows 95 or Windows NT, Macintoshes, and UNIX workstations running Solaris operating system. In the current implementation of JADIS, the simulation and interface are implemented as Java applications. Once they are converted to be designed as applets, then the entire architecture can be run using an internet browser such as Netscape on any machine.

The JADIS architecture was evaluated in the context of an aircraft maintenance problem in airbase logistics. The classes in JADIS for this application were based on a set of principles to enhance reuse, exploit natural mapping, and rapidly test different decision making strategies. The JADIS application for air craft maintenance is an interactive simulation accommodating active human interaction and has visualization capabilities.

Java was found to be a powerful language. The ability to readily move code between multiple platforms is a powerful feature. The large number of built-in classes in the language enhanced reuse. Java's capability to not only be applicable for implementing simulations but also its use in graphical user interface design made it very powerful to developing interactive simulations. Finally, the javadoc capability in Java to quickly generate hypertext source code documentation was a valuable feature.

The findings of this research contribute to the area of interactive simulations. The JADIS architecture offers a solution to the hardware and software problems encountered in interactive simulation development. Through application of the Model-View-Controller framework, simulation model development and interface design can take place concurrently thereby potentially reducing the simulation development lifecycle cost. The architecture also facilitates the study of human interaction with complex systems and the effectiveness of tailored views to interactive simulations. The airbase logistics problem studied in evaluating the JADIS architecture appears to be a ripe application area for implementing interactive simulations.

Acknowledgments

The research reported in this paper was sponsored by the United States Air Force, Logistics Research Division, Armstrong Laboratory, Wright-Patterson AFB, under contract with Research and Development Lab. We thank Capt. Todd Carrico and 2Lt. John DiPasquale at Armstrong Laboratory for their support and encouragement during this work.

References

Bell, P.C. & O'Keefe, R. M. (1987). Visual interactive simulation - History, recent developments, and major issues. *Simulation,* 49(3):109-116.

Bishop, J. L. & Balci, O. (1990). General purpose visual simulation system. *Proceedings of the 1990 Winter Simulation Conference.* 504-512.

Carrico, T., Clark, P. K., Shute, N. J., and Zahn, E. A. (1995). Integrated Model Development Environment (IMDE) multi-function aerospace support system. Technical report, *AL/HR-TR-1995-0186,* Armstrong Laboratory, WPAFB, Dayton, OH.

Carrico, T. & Clark, P. K. (1995). IMDE support for Air Force logistics. Technical report, *AL/HR-TR-1995-0187,* Armstrong Laboratory, WPAFB Dayton, OH.

Dunkler, O., Mitchell, C. M., Govindaraj, T., & Ammons, J. C. (1988). The effectiveness of supervisory control strategies in scheduling flexible manufacturing systems. *IEEE Transactions on Systems, Man, and Cybernetics,* SMC-18, 223-237.

Gobbetti, E. & Turner, R. (1991). Object-oriented design of dynamic graphics applications. *In New Trends in Animation and Visualization*, N. M. Thalmann and D. Thalmann Eds., Wiley Professional Computing, Sussex, England, 43-58.

Goldberg, A. (1990). Information models, views, and controllers. *Dr. Dobb's Journal,* July, 54-61.

Hurrion, R.D. (1980). An interactive visual simulation system for industrial management. *European Journal of Operational Research,* 5(2): 86-94.

Kirkpatrick, P.F. & Bell, P.C. (1989). Visual interactive modeling in industry: Results from a survey of visual interactive model builders. *Interfaces,* 19, 71-79.

Krasner, G. E. & Pope, S. T. (1988). A cookbook for using the model-view controller user interface paradigm in Smalltalk 80. *Journal of Object Oriented Programming,* August/September, 26-49.

Lemay, L. & Perkins, C. L. (1996). *Teach yourself Java in 21 days.* Sams.Net Publishing.

Lyu, J. & Gunasekaran, A. (1993). Developing a visual interactive simulation model for flexible manufacturing systems. *International Journal of Operations and Production Management,* 13(6): 59-67.

Narayanan, S., Bodner, D. A., Sreekanth, U., Govindaraj, T., McGinnis, L. F., & Mitchell, C. M. (1996). Research in object-oriented manufacturing simulations: An assessment of the state of the art. Submitted to *IIE Transactions*.

Paul, R. J. (1989). Visual simulation: Seeing is believing. In *Impacts of Recent Computer Advances on Operations Research*. R. Sharda, B. L. Golden, E. Wasil, O. Balci, and W. Stewart, Eds. Elsevier Science Publishing, New York, NY, 422-432.

Mathewson, S. C. (1984). The application of program generator software and its extensions to discrete event simulation modeling. *IIE Transactions*. 16(1): 3-17.

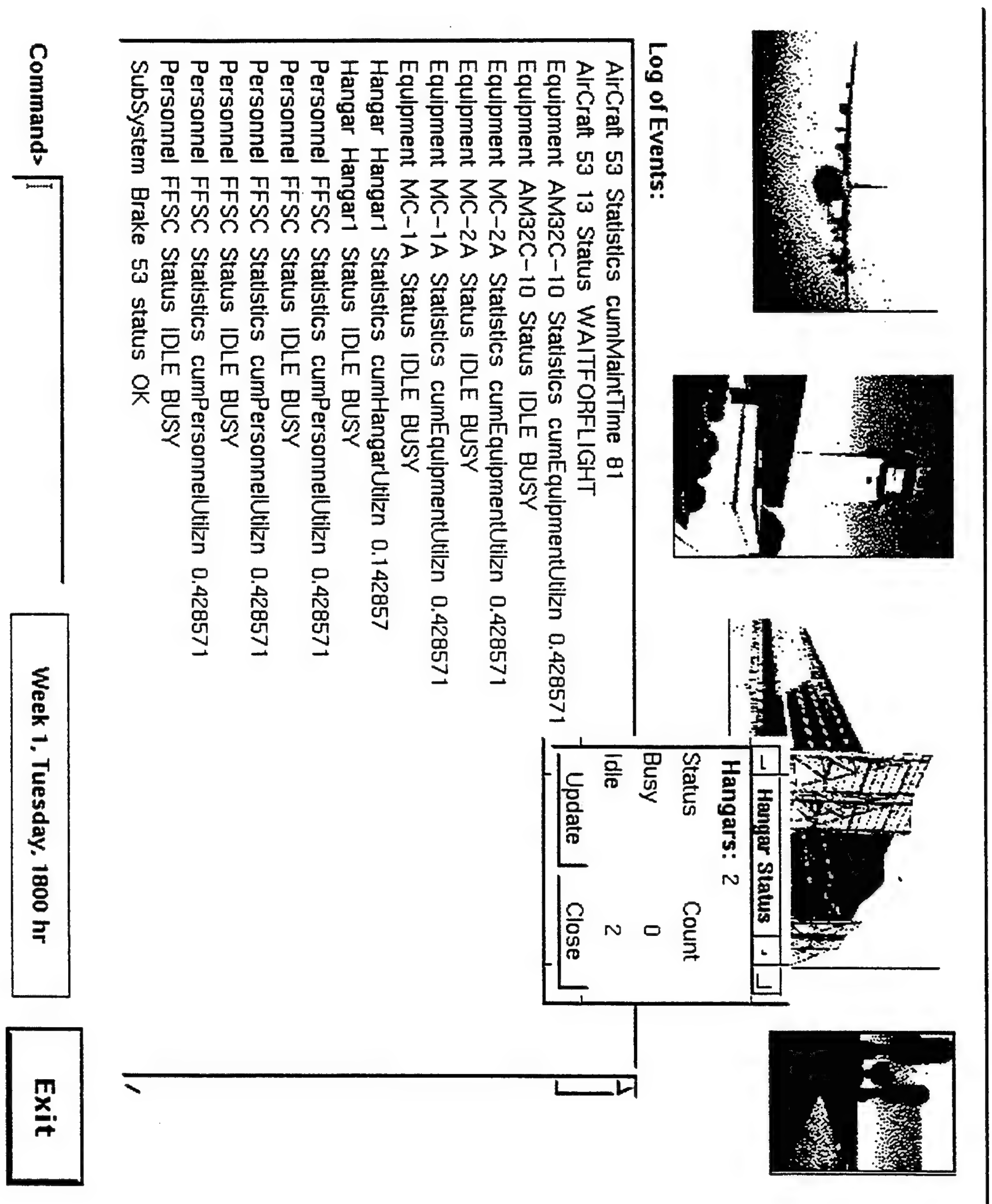
McGregor, D. R. & Randhawa S. U. (1994). ENTS: An interactive object-oriented system for discrete simulation modeling. *Journal of Object-Oriented Programming*, January, 21-29.

Popken, D. A. (1992). An object-oriented simulation environment for airbase logistics. *Simulation*. 59(5): 328-338.

Rooks, M. (1993). A user-centered paradigm for interactive simulation. *Simulation*, 60(3): 168-177.

Shan, Y-P. (1990). MoDE: A UIMS for Smalltalk. University of North Carolina, Chapel Hill, *TR90-017*, DTIC, April.

Appendix A: Screen output of the main interface.



Appendix B: Javadoc output for class `Rnormal`.

Class Rnormal

```
java.lang.Object
   |
   +-----SimBase
            |
            +----Distribution
                     |
                     +----Rnormal
```

---

public class **Rnormal**
extends Distribution

The **Rnormal Class** creates a Normal Distribution with a mean of zero and a variance of one (i.e. N(0,1)).

Version:
July 16, 1996
Author:
S. Narayanan

Constructor Index

- **Rnormal**(long)
  The overloaded *Rnormal* Constructor calls the Distribution's constructor to set the seed

Method Index

- **getNextRnormal**()
  The *getNextRandom* Method gets next random number of the standard normal distribution.

Constructors

Rnormal

```
public Rnormal(long firstSeed)
```

The overloaded *Rnormal* Constructor calls the Distribution's constructor to set the seed

Methods

- **getNextRnormal**

```
public double getNextRnormal()
```

The *getNextRandom* Method gets next random number of the standard normal distribution. It gives the n(0,1) deviate by composition method of ahrens and dieter (see Brately, Fox, and Schrage, pg. 318).

THE ABILITY TO REPRODUCE PROJECTIVE INVARIANTS OF CONICS

Christopher S. Schreiner
Graduate Assistant
Department of Psychology

Miami University
Oxford, OH 45056

Final Report for:
Summer Graduate Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

and

Armstrong Laboratory

August 1996

THE ABILITY TO REPRODUCE PROJECTIVE INVARIANTS OF CONICS

Christopher S. Schreiner
Graduate Assistant
Department of Psychology
Miami University

ABSTRACT

Two experiments were conducted to examine observers' performance on reproducing projective invariants of ellipses. Of interest was whether their performance was stable across time, and also across variations of the presentation of stimuli, such as the observer's viewpoint, the amount of movement of the ellipses, and the presence or absence of a background providing depth cues. Absolute projective invariants for pairs of conics were calculated from the observers' productions. Results indicated that performance varies from subject to subject, and that neither viewpoint, tilt, or background have a significant effect on observers' reproductions.

THE ABILITY TO REPRODUCE PROJECTIVE INVARIANTS OF CONICS

Christopher S. Schreiner

INTRODUCTION

Shape constancy and object perception are important phenomena in the field of visual perception. Much attention in the past several years has been given to the role that perceiving projective invariants plays in these aspects. This attention is at least partly due to the role invariants play in Gibson's (1950) theory of perception. Invariants are playing an ever growing role in computer vision (see Mundy and Zisserman, 1992 and Mundy and Zisserman, 1993 for applications of invariants in computer vision), but until recently little has been done by means of experiment to determine the role of invariants in human vision (Lappin and Preble, 1975).

One model concerning the concept of shape constancy depends upon the ability of the visual system to recognize projective invariants, or properties of a projected image that remain invariant over changes in viewpoint (Gibson, 1950; Biederman and Gerhardstein, 1993). A major reason that projective invariants are being investigated concerns the basic problem in perception of how distance can be seen. Distance is not preserved when light reaches the eye, but projective properties are. Gibson (1970) claimed that "we shall have to understand the mathematical invariants under optical transformation instead of the old-fashioned 'cues for depth' if we hope ever to understand space perception". The purpose of these experiments is to measure how well observers can perceive these projective invariants in conics across varying conditions, including changes in viewpoint.

EXPERIMENT ONE

There has been little research that shows that observers' estimates of invariants are stable and reliable. Experiment One investigates how robust observers' reproductions of invariants are across changes in the observer's viewpoint and the amount that the ellipses are allowed to tilt and slant within the field of vision.

METHOD

Subjects. The author and one female employee of Armstrong Laboratory (Williams Air Force Base, Arizona) participated in this experiment. Both participants had corrected to normal vision. Neither participant in this experiment took part in the other experiment reported in this paper.

Materials. The stimuli used in this experiment consisted of eighteen pairs of ellipses. There were three basic forms of ellipses, distinguished by the ratio of the minor axis to the length of the major axis. The ratios for the three ellipses were as follows: 5:1, 9:1, and 9:5. All possible combinations of the three ellipses were created. In addition, for each pair, the major axis of one of the ellipses was rotated with respect to the major axis of the other to one of three angular disparities: 15°, 45°, or 75°. Applying each of these orientations to the six pairs of ellipses created a total of eighteen distinct pairs of ellipses, which are shown in Figure 1. Each pair shared a common center. One of the ellipses was outlined in red and the other was outlined in blue.

Insert Figure 1 about here

The experiment was conducted using the Elliptical Perceptions Test software program implemented on a Silicon Graphics Iris Crimson machine.

The dependent variables were the log ratios of the two projective invariants of the standard and comparison ellipses, which will be called Invariant One and Invariant Two. The log ratio is calculated by subtracting the natural logarithm of the value of the actual invariant from the natural logarithm of the value of the estimated invariant. These invariants can be calculated from the numeric coefficients of the standard equation of each ellipse:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0.$$

There are two such equations which are used for each invariant, one for each ellipse. The coefficients are arranged in two separate matrices:

$$\underline{X} = \begin{matrix} A_1 & F_1 & E_1 \\ F_1 & B_1 & D_1 \\ E_1 & D_1 & C_1 \end{matrix} \qquad \underline{Y} = \begin{matrix} A_2 & F_2 & E_2 \\ F_2 & B_2 & D_2 \\ E_2 & D_2 & C_2 \end{matrix}$$

The determinant of each matrix is the discriminant of the equation of each conic. Invariants can be formed from these discriminants. The inverse of one matrix, the product of that inverse with the other matrix, and the trace of that product are computed. These are Invariant One and Invariant Two:

$$\mathrm{Inv}_1 = \mathrm{Trace}\,[X^{-1}\,Y]$$

$$\mathrm{Inv}_2 = \mathrm{Trace}\,[Y^{-1}\,X]$$

These invariants are absolute projective invariants of pairs of conics and are not found for individual conics, since any conic may be transformed into any other conic by a means of projection. The logarithm of these invariants has the special property of being a distance measure and therefore lies on an absolute scale of measurement (Julesz, 1971).

Procedure. The participants' task was to recreate pairs of ellipses from a standard figure of a pair of ellipses that appeared on the monitor. The standard figures appeared on the left side of the monitor with a gray and green checkerboard pattern "floor" receding into the distance dependent upon the observer's viewpoint angle. The standard figures were in continuous motion, rotating around the common center of the ellipses constrained by the tilt/slant limit. The motion occurred in random steps of five degrees every 100 msec. The comparison figures appeared on the right side of the monitor on a neutral gray background. Participants could manipulate each of the comparison ellipses by using a

mouse device to change the eccentricity of an ellipse, alter the size, or rotate an ellipse around its center. Participants could also move an ellipse in its entirety around the screen.

Participants were instructed to attempt to reconstruct the ellipses as if they were seen from face on in the picture plane; that is, they were to imagine how the ellipses would appear flat on the ground and with the observer directly over them. After recreating the ellipses, participants gave a confidence rating between 0 and 100. Participants worked at their own pace.

There were two independent variables in this experiment. One was the viewpoint angle of the observer relative to the y-axis in the three coordinate plane. The values this variable assumed were 0°, 25°, 50°, and 75°. The second independent variable was tilt/slant limit, which is the angle of maximum deviation that the two ellipses moving in tandem could vary from the ground plane. The values this variable assumed were 20°, 40°, 60°, and 80°. Each possible combination of viewpoint angle and tilt/slant limit produced 16 blocks of trials. Each block was completed twice. In each block, each of the 18 possible combinations of ellipses appeared three times, creating 54 trials per block. Participants completed a total of 1728 trials over the course of several days. All blocks and trials were presented in random order.

RESULTS AND DISCUSSION

Both log ratios of the invariants were computed and analyzed separately as a time series. As shown in Figure 2, the observers' performance differed with regard to each other. The performance of Observer Two remained relatively stable across time, while

Observer One's performance improved during the early blocks and became stable across time. However, Observer One generally performed better than Observer two, since the values of the log ratios for Observer One were generally closer to zero.

Insert Figure 2 about here

The difference in performance may also be seen in computing the correlation between the actual (ln) invariant and the estimated (ln) invariant for both observers (see Figure 3). Observer One reproduced these invariants well. For Invariant One, Observer One obtained a value of R^2 of 0.95. The slope for the regression line was relatively close to a one to one ratio ($y = 0.9532x + .0872$). For Invariant Two, Observer One obtained a value of R^2 of 0.91. The equation for the regression line was: $y = 0.8791x + 0.1826$. However, Observer Two did not perform as well, obtaining a value for R^2 for Invariant One of 0.64. The slope of the regression line did not resemble a one to one ratio ($y = 0.4966x + 0.0329$). For Invariant Two, Observer Two had a value for R^2 of 0.57 and the equation for the regression line was: $y = 0.8108x - 0.2748$.

Insert Figure 3 about here

These findings are of interest because the observers produced very different results. The results of Observer Two are difficult to interpret except to say that Observer Two was not able to reproduce the projective invariants very well. Observer One was able to adequately reproduce the invariants over time. However, Observer One did not perform as well in the first few blocks of trials. It is possible that this was simply due to familiarity with the task or the stimuli. However, those specific blocks are blocks in which the viewpoint angle was greatest (75°) and the tilt/slant limit was larger than 20°. Furthermore, this was the only viewpoint angle used in this experiment in which the ellipses appeared above the horizon line provided by the checkerboard "floor" receding into the distance. Therefore, there were multiple interpretations for these results. It may have been practice, or an extreme viewpoint angle, or else a lack of a background. Experiment Two was conducted in order to investigate whether viewpoint angle, tilt/slant limit, or absence of a background affected observers' reproductions of these invariants.

EXPERIMENT TWO

METHOD

Subjects. Seven employees of Armstrong Laboratory (Williams Air Force Base, Arizona) participated in this experiment, of whom five were men and two were women. Ages ranged from 23 to 40. All participants had normal or corrected to normal vision.

Materials. The same stimuli, equipment, and dependent variables found in Experiment One are used in this experiment. There were three independent variables in this experiment. The first independent variable was viewpoint angle, which is the angle relative to the y-axis of the three-coordinate plane from which the observer viewed the ellipses. The two values which this variable could take were 0° and 75°. The second variable was tilt/slant limit, which is the maximum angle that the ellipses, moving in tandem, could deviate from the ground plane. The values for this variable were either 20° or 80°. The third independent variable was the background used for the standard ellipses. There was either a solid background or a gray and green checkerboard background depicted in perspective providing depth cues.

Procedure. The participants' basic task was identical to that in Experiment One. Participants completed eight blocks containing 18 trials each (each ellipse appears once per block). The eight blocks were spread out over two sessions. One session consisted of four blocks. Each session lasted approximately an hour. The variables in this experiment were the viewing angle (either 0°, meaning directly over the ellipses, or 75°), the amount in degrees that the ellipses were allowed to rotate away from their standard position (either 20° or 80°), and whether or not there was a checkerboard "floor" depicted in perspective providing a depth cue. All combinations of these variables constituted the eight blocks. The order of the blocks was random for each participant.

RESULTS AND DISCUSSION

Projective invariants of each pair of ellipses were calculated. Sixteen percent of the log ratios of invariants of one observer were logs of negative numbers and therefore not real numbers. Therefore, his results were excluded from he analysis. Two 2x2x2 (Viewpoint x Tilt/Slant Limit x Background) Analyses of Variance (ANOVA) were performed separately on each set of invariant log ratios. No results of either analysis reached appropriate levels of significance. Therefore, it can be concluded that there was no effect of either viewpoint, tilt, or background on observers' reproduction of invariants.

Examining the performances of each individual separately, there is again a wide disparity as in Experiment One. The equation for each regression line measuring the actual (ln) invariant against the estimated (ln) invariant are presented in Table 1, along with the correlation coefficient R^2 for each equation. While some observers adequately reproduced the invariants, others did not. Observer's correlations of their estimates to the actual invariants ranged from .32 to .92; and while some were able to reproduce these invariants close to a one-to-one scale, others' estimations came closer to a one-to-two scale.

Insert Table 1 about here

GENERAL CONCLUSIONS

The results of these two experiments appear to indicate that some people are able to reproduce projective invariants of conics rather well while others can not. The inferior performance of some could be unfamiliarity with the task used in these experiments. Regardless of how well their performance is, it does not appear to be affected by the observer's viewpoint, by the amount of tilt of the shapes, or by the presence or absence of a background. This finding in Experiment Two leads one to conclude that the performance in the first few blocks of Observer One in Experiment One was caused by familiarity with the task. It would appear then, that observers' performance at perceiving and reproducing these invariants does become stable across all conditions.

These results do not necessarily move towards answering whether there is a psychological mechanism in the visual system that allows for the computation of these invariants or whether these invariants may provide an explanation for shape constancy. Instead, they should be used as a measurement of shape constancy. Projective properties such as these can be used to demonstrate how well we see, to be applied in different experimental conditions that may ultimately provide an answer as to how we see, but they are not an explanation in and of themselves.

REFERENCES

Biederman, I. & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 1162-1182.

Gibson, J. J. (1950). *The Perception of the Visual World.* Cambridge, MA: Riverside Press.

Gibson, J. J. (1970). On theories for visual space perception. *Scandinavian Journal of Psychology, 11,* 75-79.

Julesz, B. (1971). *Foundations of Cyclopean Perception.* Chicago: University of Chicago Press.

Lappin, J. S. & Preble, L. D. (1975). A demonstration of shape constancy. *Perception & Psychophysics, 17,* 439-444.

Mundy, J. L. & Zisserman, A. (Eds.). (1992). *Geometric Invariance in Computer Vision.* Cambridge, MA: MIT Press.

Mundy, J. L. & Zisserman, A. (Eds.). (1993). *Applications of Invariance in Computer Vision II.* Ponta-Delgada, Azores.

CAPTIONS

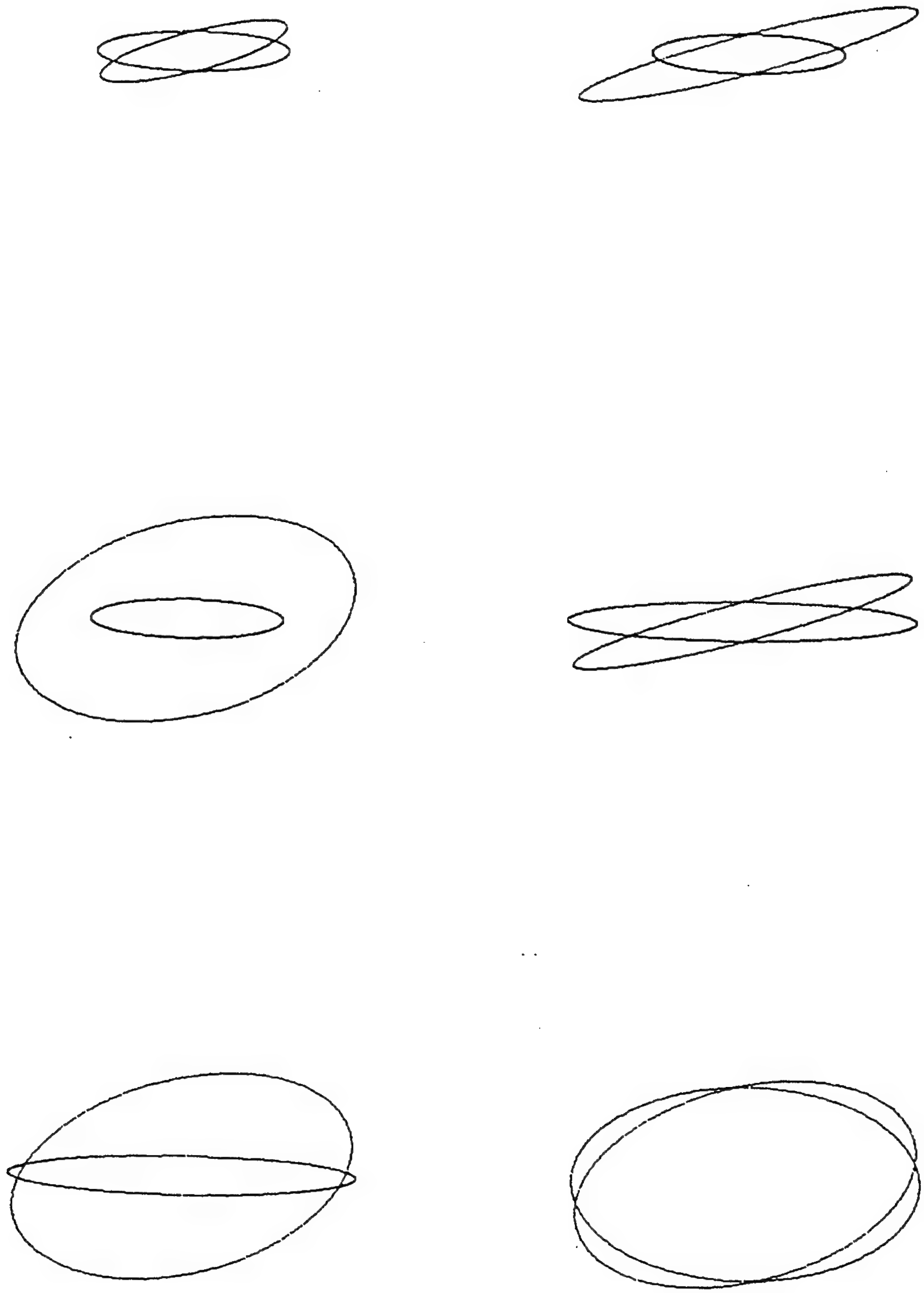
Figure 1. The eighteen pairs of ellipses used in both Experiment One and Experiment Two.

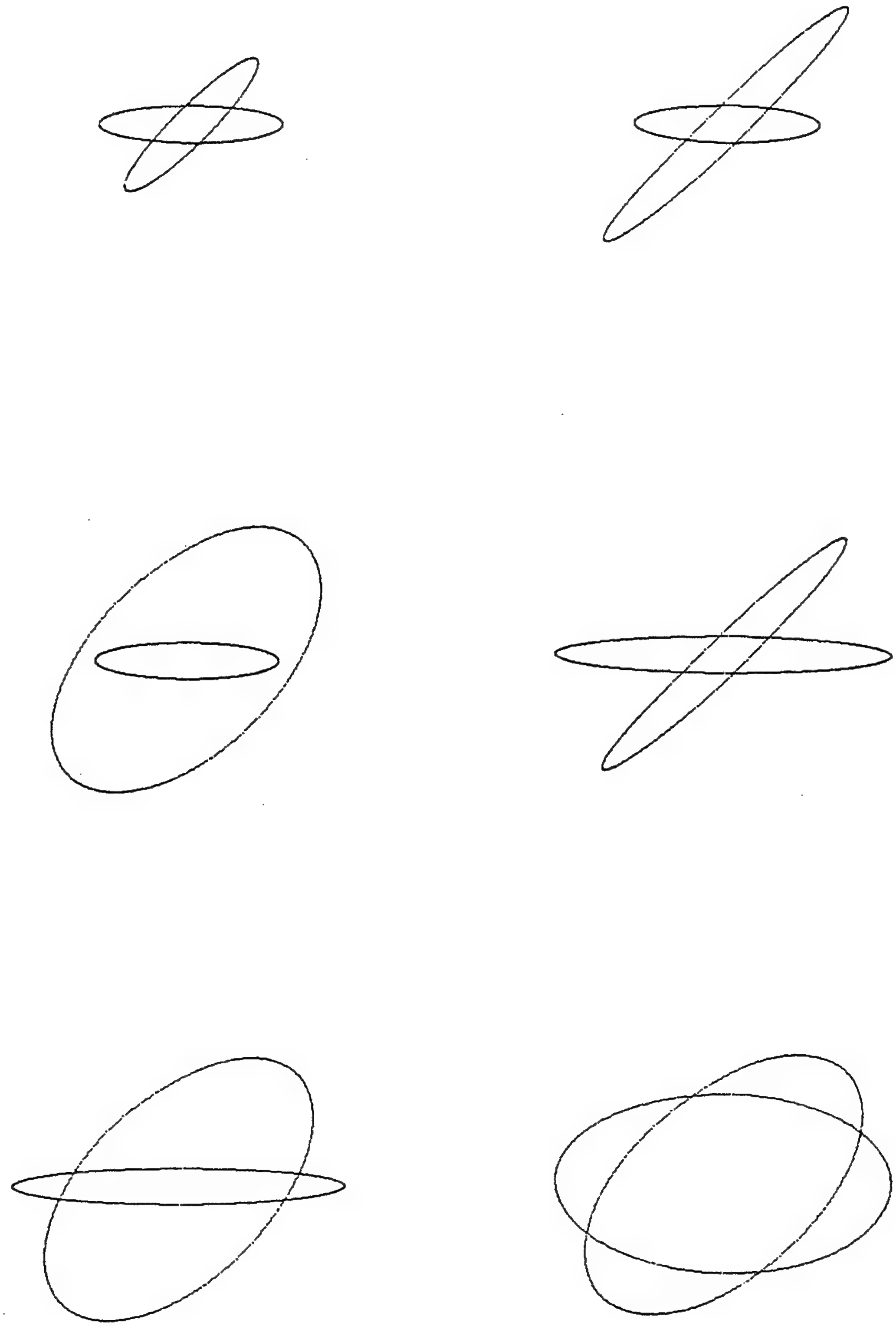
Figure 2. Time series analysis showing the moving average of the log ratios of invariants for Observer One and Observer Two in Experiment One..

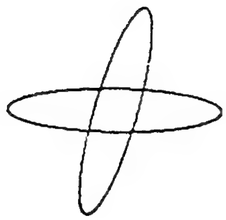
Figure 3. Mean values of Observer One's and Observer Two's estimates of two invariants versus the actual values of the invariants in Experiment One.

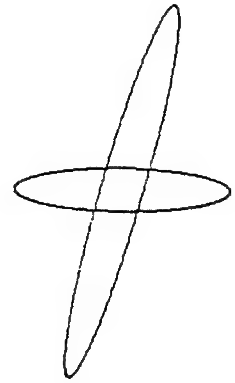
Table 1. Regression coefficients and equations for the six observers in Experiment Two.

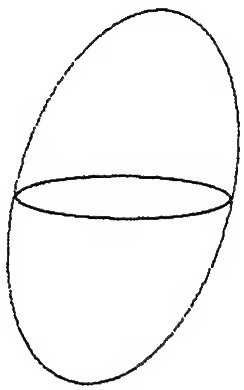
Figure 1.

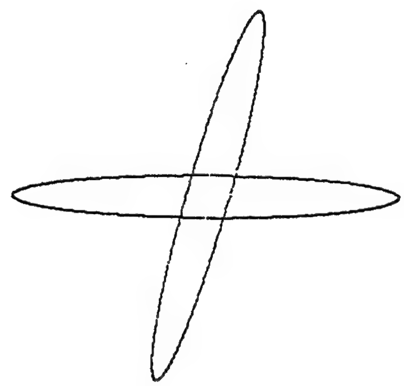


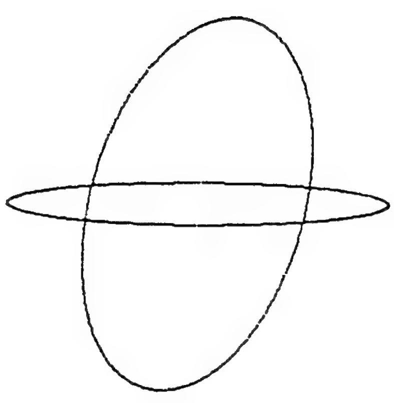












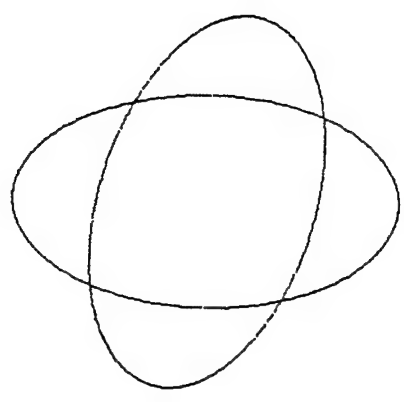
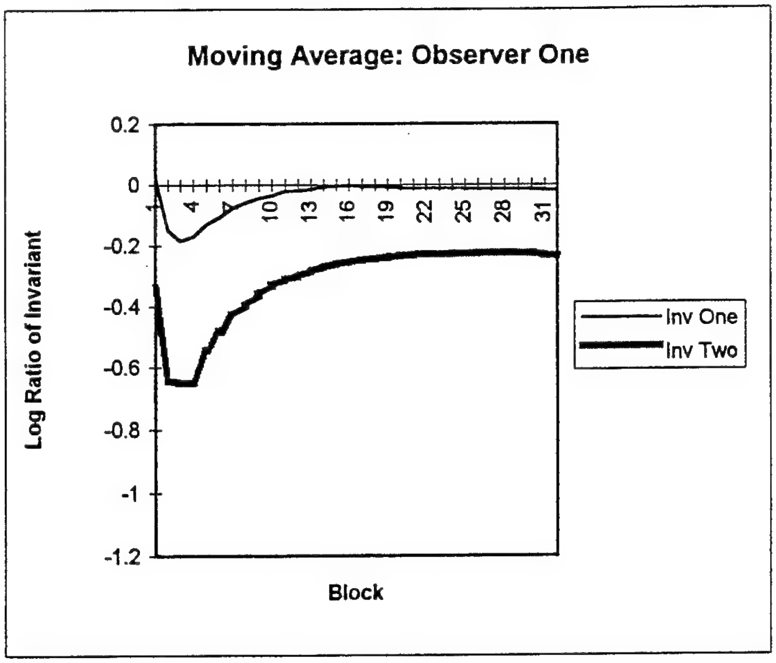
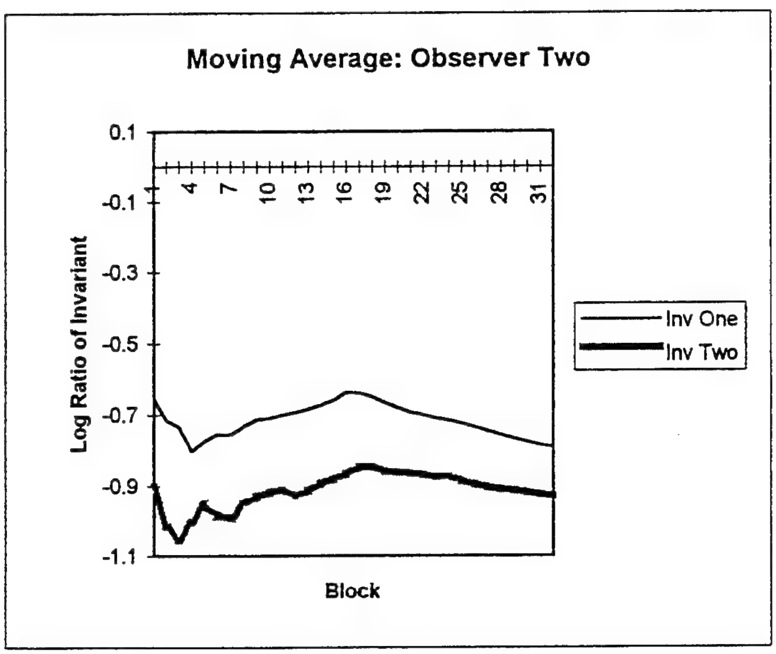


Figure 2.

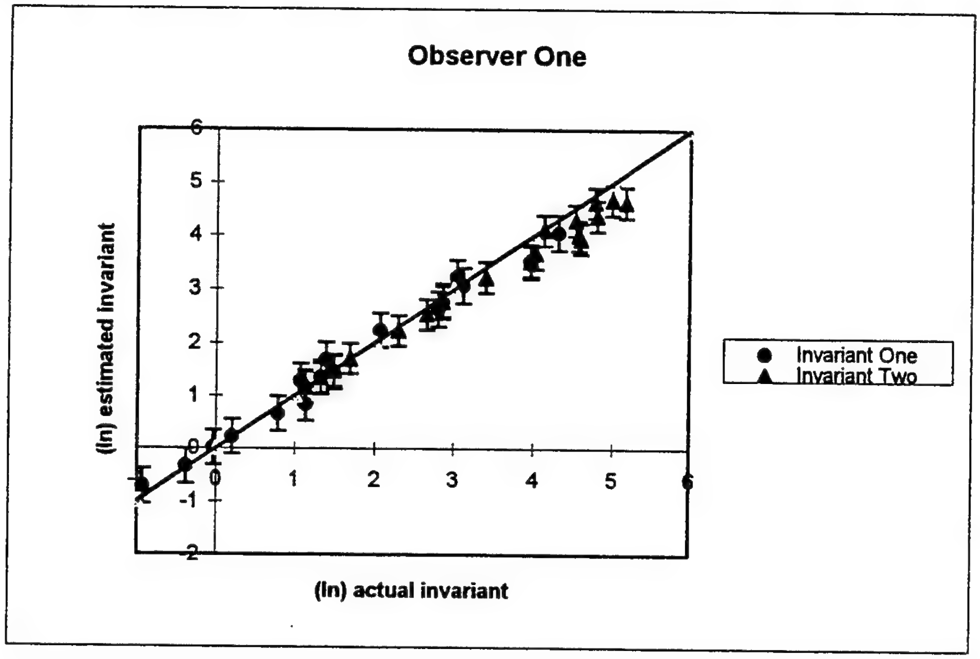


Moving Average: Observer One
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
1
4
7
10
13
16
19
22
25
28
31
Log Ratio of Invariant
Block
Inv One
Inv Two

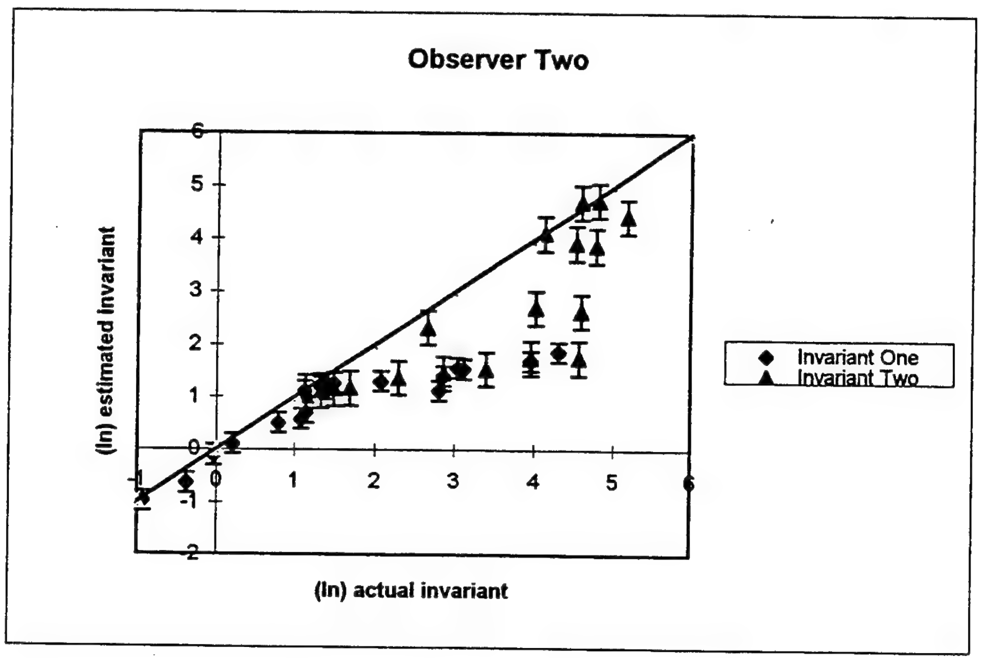


Moving Average: Observer Two
0.1
-0.1
-0.3
-0.5
-0.7
-0.9
-1.1
1
4
7
10
13
16
19
22
25
28
31
Log Ratio of Invariant
Block
Inv One
Inv Two

Figure 3.



Observer One
(ln) estimated invariant
(ln) actual invariant
6
5
4
3
2
1
0
-1
-2
0
1
2
3
4
5
6
Invariant One
Invariant Two



Observer Two
(ln) estimated invariant
(ln) actual invariant
6
5
4
3
2
1
0
-1
-2
0
1
2
3
4
5
6
Invariant One
Invariant Two

Table 1.

Observer JD		
Invariant One	$R^2 = 0.61$	$y = 0.83\,x + 0.23$
Invariant Two	$R^2 = 0.51$	$y = 0.88\,x + 0.03$
Observer PG		
Invariant One	$R^2 = 0.42$	$y = 0.58\,x + 0.22$
Invariant Two	$R^2 = 0.35$	$y = 0.55\,x + 0.24$
Observer GA		
Invariant One	$R^2 = 0.90$	$y = 0.90\,x - 0.04$
Invariant Two	$R^2 = 0.67$	$y = 0.80\,x + 0.04$
Observer BS		
Invariant One	$R^2 = 0.81$	$y = 0.87\,x - 0.04$
Invariant Two	$R^2 = 0.72$	$y = 1.01\,x - 0.43$
Observer RT		
Invariant One	$R^2 = 0.67$	$y = 0.75\,x + 0.30$
Invariant Two	$R^2 = 0.32$	$y = 0.66\,x + 0.24$
Observer JC		
Invariant One	$R^2 = 0.92$	$y = 0.90\,x + 0.11$
Invariant Two	$R^2 = 0.84$	$y = 0.89\,x - 0.09$

ARITHMETIC EFFECTS ON AIMING PERFORMANCE IN COORDINATION: SEQUENTIAL POSITION EFFECTS

Jacqueline C. Shin
Department of Psychology

Pennsylvania State University
University Park, PA 16802

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base. DC

and

Armstrong Laboratory

ARITHMETIC EFFECTS ON AIMING PERFORMANCE IN COORDINATION: SEQUENTIAL POSITION EFFECTS

Jacqueline C. Shin
Department of Psychology
Pennsylvania State University

ABSTRACT

Most tasks require coordination of cognitive functioning and perceptual-motor control. For example, an airplane pilot must make high-level decisions about navigation and also reach for and manipulate knobs in the cockpit. The demands of the perceptual-motor tasks can in principle affect achievement of the cognitive tasks, and vice-versa. Remarkably, despite the large amount of research that has been done on isolated perceptual-motor tasks, on isolated cognitive tasks, and on concurrent tasks (tasks that must be performed at the same time), there has been virtually no study of tasks whose defining feature is satisfaction of perceptual-motor goals in the service of cognitive goals. The current study was part of a series of studies that explores how arithmetic processes affect continuous aiming performance when aiming is performed in the service of sequential arithmetic. My goal was to understand a specific pattern of results thought to reflect arithmetic influences on aiming performance, the sequential position effect, characterized by longer aiming times at later steps in an aiming-arithmetic sequence. The results of the study strongly suggest that the sequential position effect reflected active working-memory management processes of arithmetic information, perhaps those involved in evaluating and reconstructing arithmetic calculation results or instantiating a plan for arithmetic calculations to take place.

ARITHMETIC EFFECTS ON AIMING PERFORMANCE IN COORDINATION: SEQUENTIAL POSITION EFFECTS

Jacqueline C. Shin

INTRODUCTION

This study is part of series of studies that investigates how cognitive and perceptual-motor activities are coordinated in a goal-directed situation. This line of research uses the *nested aiming-arithmetic* task, in which aiming is performed in the service of mental arithmetic. In this task, participants moved a cursor back and forth between two targets on a computer monitor using a computer mouse or joystick to request an operator-operand pair at each step to perform sequence of arithmetic calculations. Previous experiments with this task showed that aiming time (step time, i.e., dwell time plus movement time) varied as a function of step number. Specifically, step time was shorter at earlier than later steps. I will call this pattern of times the *sequential position* effect. In the current study, I explored three hypotheses about the source of the sequential position effect -- what I called the *calculation-postponement* hypothesis, the *memory-load* hypothesis, and the *checking* hypothesis.

According to the calculation-postponement hypothesis, calculations may have been put off until later in the sequence, so that the sequential position effect reflected the likelihood that arithmetic calculations were performed at each step. Assuming that it takes a measurable amount of time to perform arithmetic calculations, the sequential position effect would have then reflected the growing likelihood of the need to perform calculations as a function of step number. According to the memory-load hypothesis, the sequential position effect reflected the number of arithmetic elements ("+" and "-" signs and numerals) held in working memory. Provided that information from earlier steps was not erased from working memory, processing would be slowed to the extent that the residual memory elements would interfere with ongoing calculations. The third hypothesis I considered about the source of the sequential position effect was the checking hypothesis. According to this hypothesis, the sequential position effect reflected an increase in active memory processing with step number. Because the arithmetic sequence was fairly long, partici-

pants may have performed working memory processes in order to keep track of their place in the entire sequence. These processes may have taken more time at later steps, because one's representation of the place in the entire sequence may have become more complex at the later steps. Moreover, because the task instructions emphasized arithmetic accuracy, participants may have checked past calculations when confidence about previous calculations was low. These processes may have slowed processing at later steps, more so the later the step number.

To test the above mentioned hypotheses, I manipulated the probability that arithmetic calculations could occur at each step. One group of participants (the *calculation-based movement* group) was supposed to move the cursor out of each target through a specific half of the target based on the even-odd status of the arithmetic result at each step. I also included two control groups who did not have to move the cursor out of the target based on the arithmetic calculation. Consider what the three hypotheses predict about performance of the calculation-based movement group versus the control groups. According to the calculation-postponement hypothesis, the sequential position effect should disappear in the calculation-based movement group but should not disappear in the control groups. The reason for this prediction is that, according to the calculation-postponement hypothesis, the sequential position effect is due to optional postponement of calculation, and when calculation postponement is no longer optional (in the calculation-based movement group), the effect of step number (the sequential position effect) should disappear. According to the memory-load hypothesis, all groups should show the sequential position effect, because as step number increases, the amount of information held in working memory should grow, in which case performance time (both dwell time and movement time) should increase with step number. Finally, according to the checking hypothesis, the sequential position effect should be found in the dwell times but not in the movement times of the calculation-based movement group, whereas the sequential position effect should be found in the dwell times and the movement times of the control groups. The reason for this prediction is that the calculation-based movement group must do all calculations before moving (i.e., during dwell), so the effects of checking previous calculations should be localized in the dwell times of these subjects. By contrast, the control

groups may perform or check previous calculations any time they wish, so both their dwell times and movement times could grow with step number.

A second goal of the study was to investigate individual differences in coordination performance. Specifically, the correlation between the nested aiming-arithmetic task performance and tests of time estimation, quantitative working memory, and figural induction abilities was explored. However, the procedures and results related to this second part of the study are not reported here, but will appear in a later more complete exposition.

METHODS

The study consisted of two parts. In the first part of the study, participants performed cognitive and perceptual-motor coordination tasks described below. In the second part, participants took a test battery measuring time estimation (a variant of the time estimation task described in Rosenbaum and Patashnik,1980), quantitative working memory (a variant of the quantitative working memory task described in Kyllonen and Christal, 1990), and figural induction abilities (a task similar to the Raven's Progressive Matrix test described in Embretson,1995; see Chaiken, 1994), which are not detailed any further here.

Participants

One-hundred eighty-two Air Force basic trainees participated in the experiment at Lackland Air Force Base in groups of 31 each in single sessions that lasted approximately three hours.

Procedure

Each participant performed three tasks -- nested aiming-arithmetic, aiming-only, and arithmetic-only tasks -- in one of three movement conditions.

Nested aiming-arithmetic task. Participants were supposed to add or subtract 5 numbers, based on a sequential display of signs ("+" and "-") and numbers (integers). As shown in Figure 1, four target circles were displayed in a diamond, and boxes were displayed next to the horizontal targets. At the beginning of each trial, the cursor was centered on the screen. When the partici-

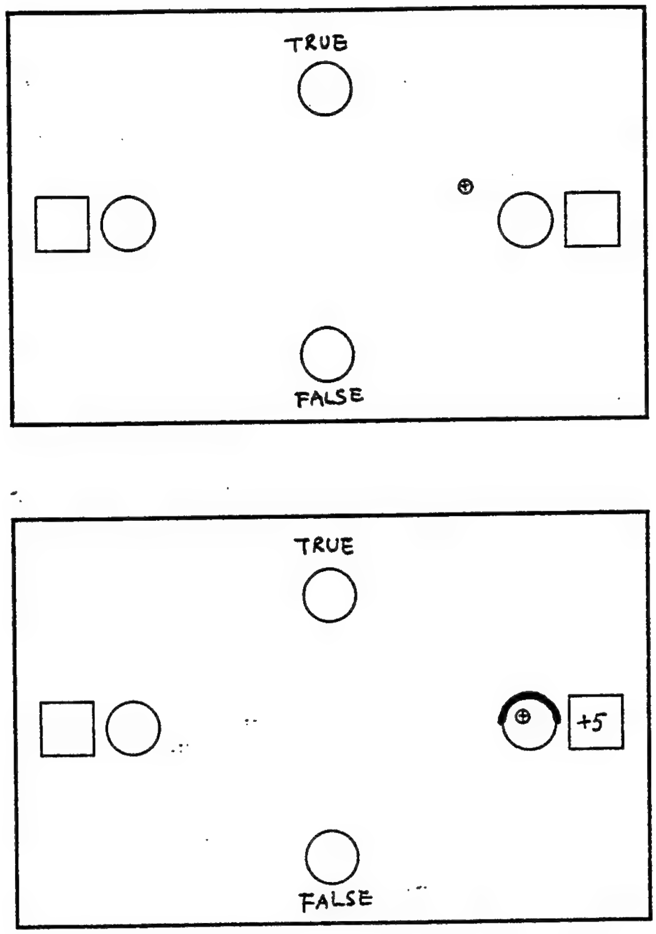


Figure 1. Experimental display for the calculation-based movement and visual cue-based movement conditions. When the participant moved the cursor into a target, the top or bottom half of the target became thickened, and when the participant moved the cursor out of the target, the thickened half of the target returned to normal thickness. In the calculation-based movement condition, the participant was supposed to move the cursor out of the thickened half if the intermediate result at that step was even or out of the opposite half if the intermediate result was odd. In the visual cue-based movement condition, the participant was always supposed to move the cursor out of the thickened half. In the no-decision condition, neither half of the target became thickened, and the participant could move the cursor out of either the top or bottom half of the target.

pant pressed a joystick button, an asterisk appeared in the left or right target. This was a signal for the participant to move the cursor to that target with a joystick. Then the participant moved the cursor back and forth between the left and right targets for a total of seven steps. The location of the cursor was isomorphically mapped to position of the joystick, and the cursor was visible throughout the trial. Each time the participant moved the cursor into a target, arithmetic stimuli were displayed in the adjacent box. As shown in Table 1, at the first five steps, a sign-numeral pair (e.g., "+5") was presented. At the sixth step, an equals sign ("=") was presented, and at the seventh step, a numeral was presented. When the participant moved the cursor out of the target, the arithmetic stimuli disappeared; when the participant moved the cursor out of and back into a given target, the stimuli did not reappear. After the seventh step, the participant was supposed to move the cursor into the top ("true") or bottom ("false") target to indicate whether the last number displayed matched the actual running total of the arithmetic sequence. This terminated the trial.

Table 1

Example of step-by-step display of arithmetic stimuli presented in squares adjacent to left and right targets on a typical trial

	Displayed location	
Step	Left	Right
1	+4	
2		+5
3	- 3	
4		+2
5	-4	
6		=
7	4	

Aiming-only task. The display of targets and arithmetic stimuli were identical to the nested aiming-arithmetic task. As in the nested aiming-arithmetic task, participants controlled the sequential display of stimuli by moving a cursor back and forth between the left and right targets. However, in the aiming-only task, instead of performing arithmetic, participants performed a simple number-monitoring task, namely, keeping track of the last number presented with a "+" sign. Plus ("+") and minus ("-") signs were presented under the constraints that in each block of trials (1) there was at least one minus sign; and (2) the last plus sign in the sequence was presented on the second, third, fourth, and fifth step an equal number of times in each block of trials. After a numeral was displayed at the seventh step, participants indicated whether the that number matched the last number presented with a "+" sign by moving the cursor into the top ("true") or bottom ("false") target.

Arithmetic-only task. The display of targets and arithmetic stimuli were identical to the nested aiming-arithmetic and aiming-only tasks. As in the nested aiming-arithmetic task, participants were supposed to add or subtract five numbers as displayed. However, instead of moving the cursor in a continuous fashion, the participant slightly displaced ("budged") the joystick in the appropriate direction. When the participant displaced the joystick from center position by a threshold amount in the appropriate direction, the cursor and arithmetic stimuli disappeared. When the participant centered the joystick, the cursor and arithmetic stimuli for the next step appeared. After the last number was displayed at the seventh step, the participant indicated whether that number matched the arithmetic total by moving the cursor to the top ("true") or bottom ("false") target.

The three tasks were designed to maximize both aiming and arithmetic demands in the nested aiming-arithmetic task, maximize aiming but minimize arithmetic demands in the aiming-only task, and minimize aiming but maximize arithmetic demands in the arithmetic-only task.

Movement Conditions. Participants were randomly assigned to three conditions that differed in how participants were supposed to move the cursor out of the targets -- the *calculation-based movement, visual cue-based movement,* and *no-decision* conditions.

In the calculation-based movement condition, each time the participant moved the cursor into a target at steps 1 through 6, the top or bottom half of the target was thickened unpredictably (see Figure 1). Participants moved the cursor out of the thick or thin half depending on the even-odd status of the arithmetic or number monitoring result at that step. If the result was even, they were supposed to move the cursor out of the thick half, and if the result was odd, they were supposed to move the cursor out of the thin half. When the participant moved the cursor out of a target, the target returned to normal thickness. Each step was scored with respect to whether the participant moved the cursor out through the appropriate half of the target. In the visual cue-based movement condition, the target became thickened in the same manner as in the calculation-based movement condition. However, participants were always supposed to move the cursor out of the thick half of the target at each step regardless of the even-odd status of the intermediate result. In the no-decision condition, the thickness of the target remained normal throughout the trial, and participants could move the cursor out of any half of the target desired. In this condition, all steps were scored as correct with respect to target exiting.

Feedback. Feedback was given to each participant at the end of each trial and block of practice. At the end of each trial, feedback was given about whether the "true"/"false" response was correct and the number of appropriately exited targets. For trials correct on both these aspects of accuracy, total time was also displayed. At the end of each block, the number of accurate trials in that block and the average total time for these trials were displayed.

Design

The study had a 3 (task) x 3 (movement condition) mixed factorial design, in which task was manipulated within subjects, and movement condition was manipulated between subjects. The tasks were presented in four cycles. Twenty four trials were presented for each task per cycle. The order of three tasks was counterbalanced among participants, such that the first two cycles and the last two cycles were presented in opposite orders.

During each trial, the computer logged the x, y position of the cursor on the computer monitor at the computer monitor refreshment rate of approximately 60 Hz. This was used to compute the step, dwell, movement times for each step. The computer also logged trial accuracy.

RESULTS AND DISCUSSION

To focus on coordination in skilled performance, I only report the results of the last quarter of practice. Step time, dwell time, and movement time were analyzed only for steps 2 through 5. Of the 182 participants that performed the coordination task, results for only 169 participants are reported because of incomplete data collection and missing cells. Also excluded from the analyses were trials that lasted longer than 60 s per step. The analyses focused on both accuracy and time measures. Step time, dwell time, and movement time were analyzed only for accurate trials. Also, trials on which participants moved the cursor to a true or false target at the wrong point in the sequence were excluded from the analyses.

Accuracy. A trial was considered accurate only if the "true"-"false" response was correct and all six targets were exited through the appropriate half. Table 2 shows the accuracy averaged over participants as a function of movement condition and task.

Table 2

Accuracy as a function of movement condition and task

	Task		
Movement Condition	Nested	Aiming	Arithmetic
Calculation-Based	.73 (.02)	.72 (.03)	.81 (.02)
Visual Cue-Based	.67 (.03)	.54 (.04)	.70 (.03)
No-Decision	.82 (.02)	.85 (.02)	.86 (.02)

*Mean (Standard Error)

A condition x task analysis of variance revealed that accuracy was highest in the no-decision condition, lowest in the calculation-based movement condition, and intermediate in the visual cue-based movement condition, $F(2,166) = 21.89$, $MSe = .08$, $p < .0001$, Newman-Keuls $ps < .01$. The main effect of task was significant, $F(2,332) = 14.71$, $MSe = .02$, $p < .0001$, such that accuracy in the arithmetic-only task was higher than in the nested aiming-arithmetic and aiming-only tasks, Newman-Keuls $ps < .001$, which did not differ from each other. The condition x task interaction was also significant, $F(2,332) = 6.18$, $MSe = .02$, $p < .001$, such that the absolute difference in accuracy between tasks varied depending on the condition.

Step Time. Step time at step n was defined as the time from when the participant moved the cursor into the nth target to when the participant moved the cursor into the $(n+1)$th target. The means of individual median step times are plotted as a function of step number and task for each condition in Figure 2.

As expected, performance in the computationally more complex tasks was slower than the simpler tasks. A condition x task x step analysis of variance revealed that step time was longest in the calculation-based movement condition, shortest in the no-decision condition, and intermediate in the visual cue-based movement condition, $F(2,166) = 110.74$, $MSe = 2223318$, $p < .0001$, Newman-Keuls $ps < .001$. Aiming was also slow when performed together with arithmetic processes relative to when performed alone. Step time was longest in the nested aiming-arithmetic task, shortest in the aiming task, and intermediate in the arithmetic-only task, $F(2,332) = 276.74$, $MSe = 461983$, $p < .0001$. A condition x task interaction, $F(4,332) = 17.68$, $MSe = 461983$, $p < .0001$, suggested that this pattern held only in the calculation-based movement and visual cue-based movement conditions as confirmed by task x step analyses of variance conducted for each condition. However, step time was not significantly different for the nested aiming-arithmetic and arithmetic tasks in the no-decision condition, suggesting that aiming time (step time) in the nested aiming-arithmetic task is slowed down ("stretched") to match the time to perform arithmetic during each step.

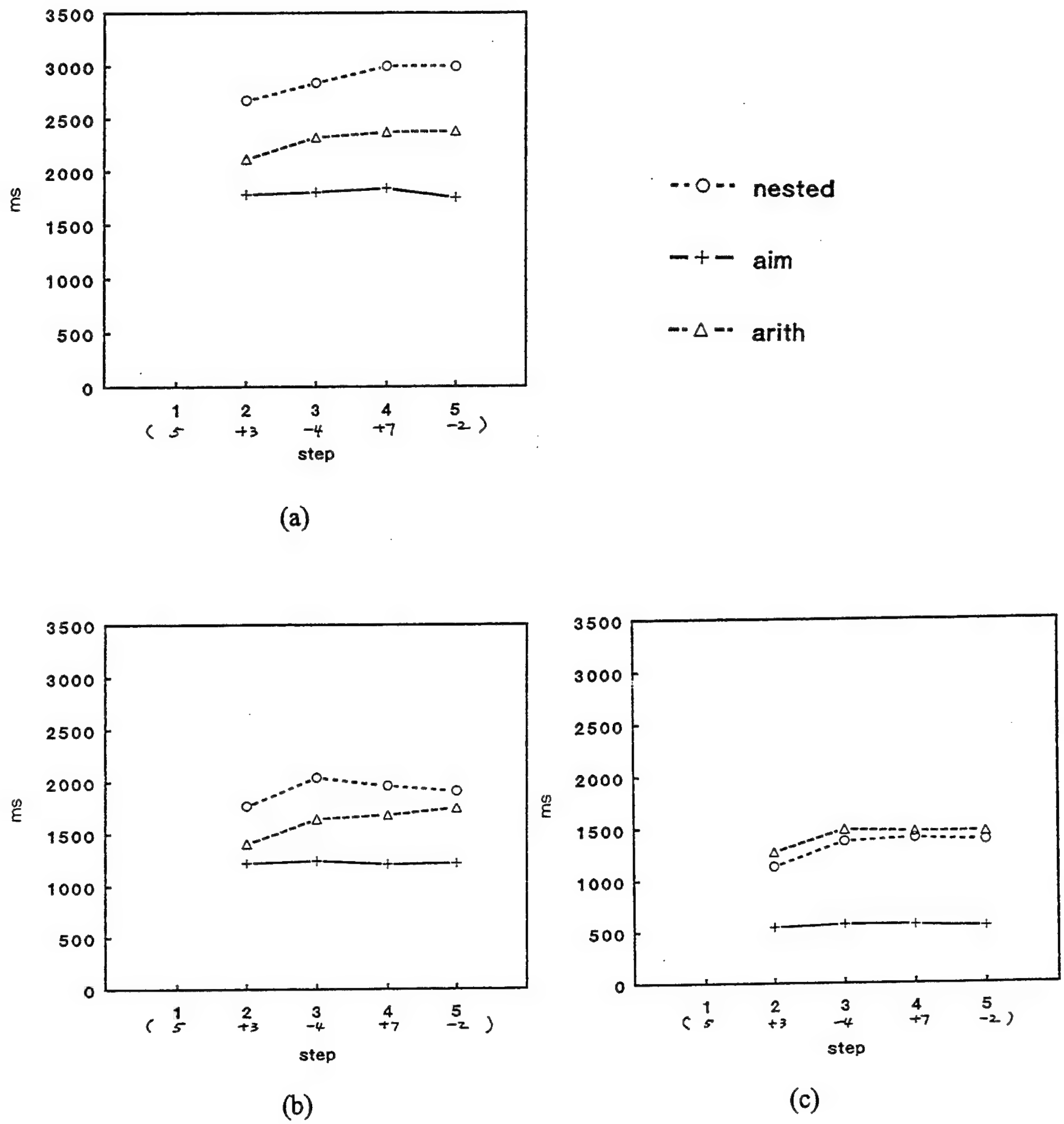


Figure 2. Step time for individual steps as a function of task and step number for (a) the calculation-based movement condition, (b) the visual cue-based movement condition, and (c) no-decision condition for the last quarter of practice.

In all conditions, step time was significantly shorter for step 2 than for steps 3 through 5, Newman-Keuls *ps* $< .01$. That sequential position effects were found in all conditions including the calculation-based movement condition, *ps* $< .0001$, argues against the postponement hypothesis but is consistent with the memory-load and checking hypotheses. The magnitude of the sequential position effect differed between tasks, as indicated by a significant task x step interaction, $F(6,996) = 13.85$, $MSe = 57661$, $p < .0001$.

Dwell time. Dwell time at step *n* was defined as the time from when the participant moved the cursor into the *n*th target for the first time to when the participant moved the cursor out of that target for the first time. The means of individual median dwell times are plotted as a function of step number and task for each condition in Figure 3.

Congruent with the results of step time, dwell time was slower in the computationally more complex tasks than the simpler tasks. A condition x task x step analysis of variance revealed that dwell time was longest in the calculation-based movement condition, shortest in the no-decision condition, and intermediate in the visual cue-based movement condition, $F(2,166) = 141.01$, $MSe = 1558737$, $p < .0001$, Newman-Keuls *ps* $< .001$. Also, dwell time in the nested task was slower than the aiming task, suggesting that arithmetic slows down dwell time. Dwell time was longest in the arithmetic-only task, shortest in the aiming-only task, and intermediate in the nested aiming-arithmetic task, $F(2,332) = 274.47$, $MSe = 422650$, $p < .0001$, Newman-Keuls *ps* $< .001$. However, a condition x task interaction, $F(4,332) = 16.65$, $MSe = 422650$, $p < .0001$, suggested that this pattern held only in the visual cue-based movement and no-decision conditions. In the calculation-based movement condition, the difference in dwell time between nested aiming-arithmetic and arithmetic-only tasks was not reliable, suggesting that arithmetic calculations affected the dwell phase to a similar extent in these tasks.

Again, sequential position effects were found in dwell time in all conditions, *ps* $< .0001$, such that dwell time was significantly shorter for step 2 than for steps 3 through 5, Newman-Keuls $p < .001$, which did not differ significantly from one another. This result argues against the calculation-postponement hypothesis, but is consistent with the memory-load and checking hypotheses.

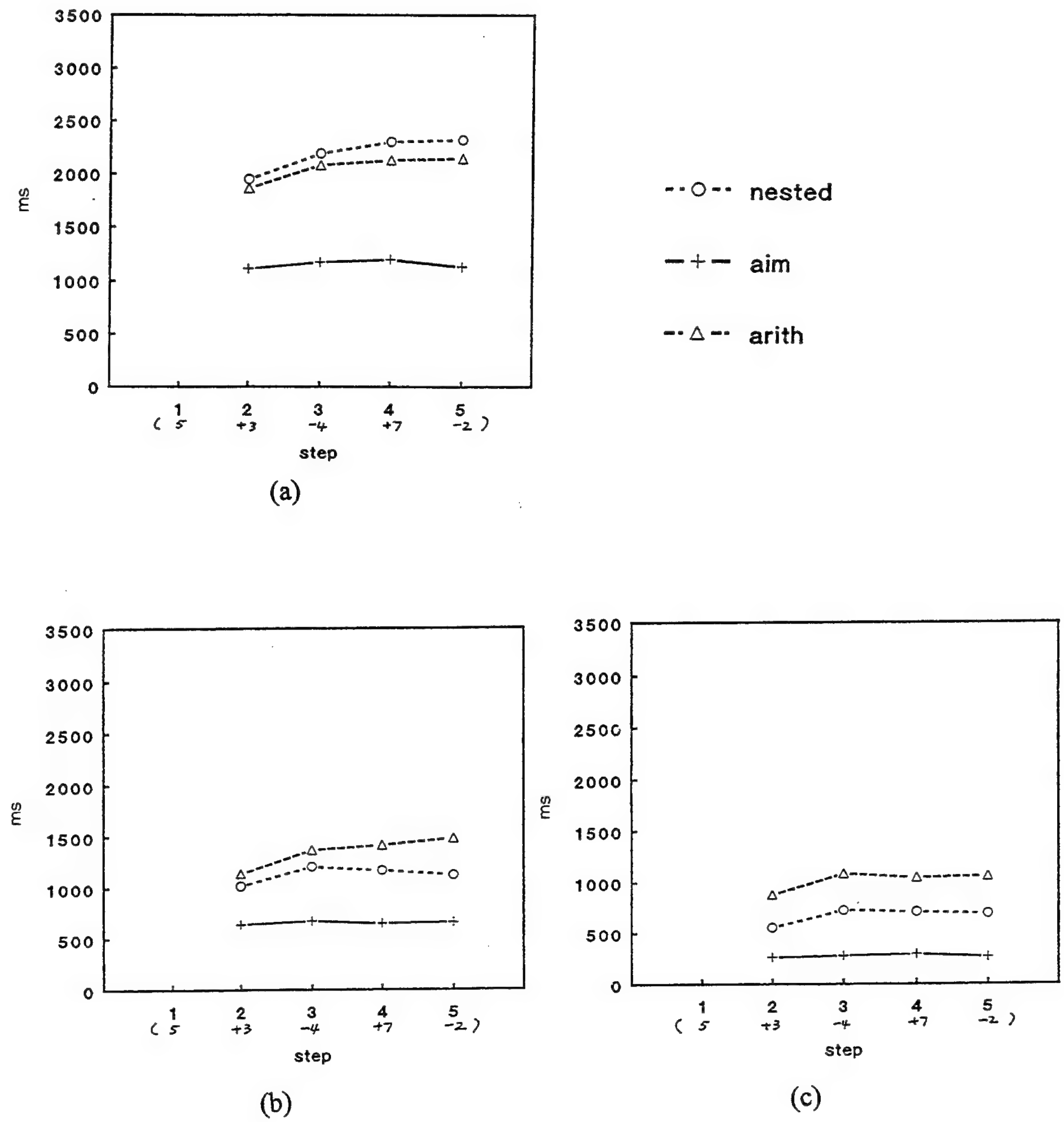


Figure 3. Dwell time for individual steps as a function of task and step number for (a) the calculation-based movement condition, (b) the visual cue-based movement condition, and (c) no-decision condition for the last quarter of practice.

As indicated by a significant task x step interaction, $F(6,996) = 12.30$, $MSe = 47741$, $p < .0001$, the sequential position effect was greater for the nested aiming-arithmetic and arithmetic-only tasks than for the aiming-only task, suggesting that the sequential position effect reflects arithmetic processes.

Movement time. Movement time at step n was defined as the time between when the participant moved the cursor out of the $(n\text{-}1)$th target to when the participant moved the cursor into the nth target. The means of individual median dwell times are plotted as a function of step number and task for each condition in Figure 4.

In contrast to step and dwell times, depending on the condition, movement time was not necessarily longer in the conditions that were computationally more complex. In a condition x task x step analysis of variance, the main effect of condition was significant, $F(2,166) = 17.77$, $MSe = 193379$, $p < .0001$, such that movement time was shorter in the no-decision condition than in the visual cue-based and calculation-based conditions, Newman-Keuls ps $< .001$. However, movement time in the visual cue-based and calculation-based conditions did not differ from each other. Movement time was longest in the nested aiming-arithmetic task, shortest in the arithmetic-only task, and intermediate in the aiming-only task, $F(2,332) = 176.48$, $MSe = 93099$, $p < .0001$, Newman-Keuls ps $< .001$. However, a condition x task interaction, $F(4,332) = 42.78$, $MSe = 93099$, $p < .0001$ suggested that this pattern held only in the visual cue-based condition as confirmed in a task x step analysis of variance in this condition, $F(2,116) = 241.17$, $MSe = 4619$, $p < .0001$, Newman-Keuls ps $< .01$. In the calculation-based condition, movement times in the nested aiming-arithmetic and aiming-only tasks did not differ from each other significantly. This suggested that arithmetic processes did not take place during movement in the nested aiming-arithmetic task. These times were longer in these tasks than in the arithmetic-only task, Newman-Keuls ps $< .01$. A different patterns of results was found in the other conditions. Arithmetic processes appear to have taken place during movement in the visual cue-based movement and no-decision conditions; movement time was longer in the nested arithmetic than in the aiming-only tasks in these conditions, Newman-Keuls ps $< .001$.

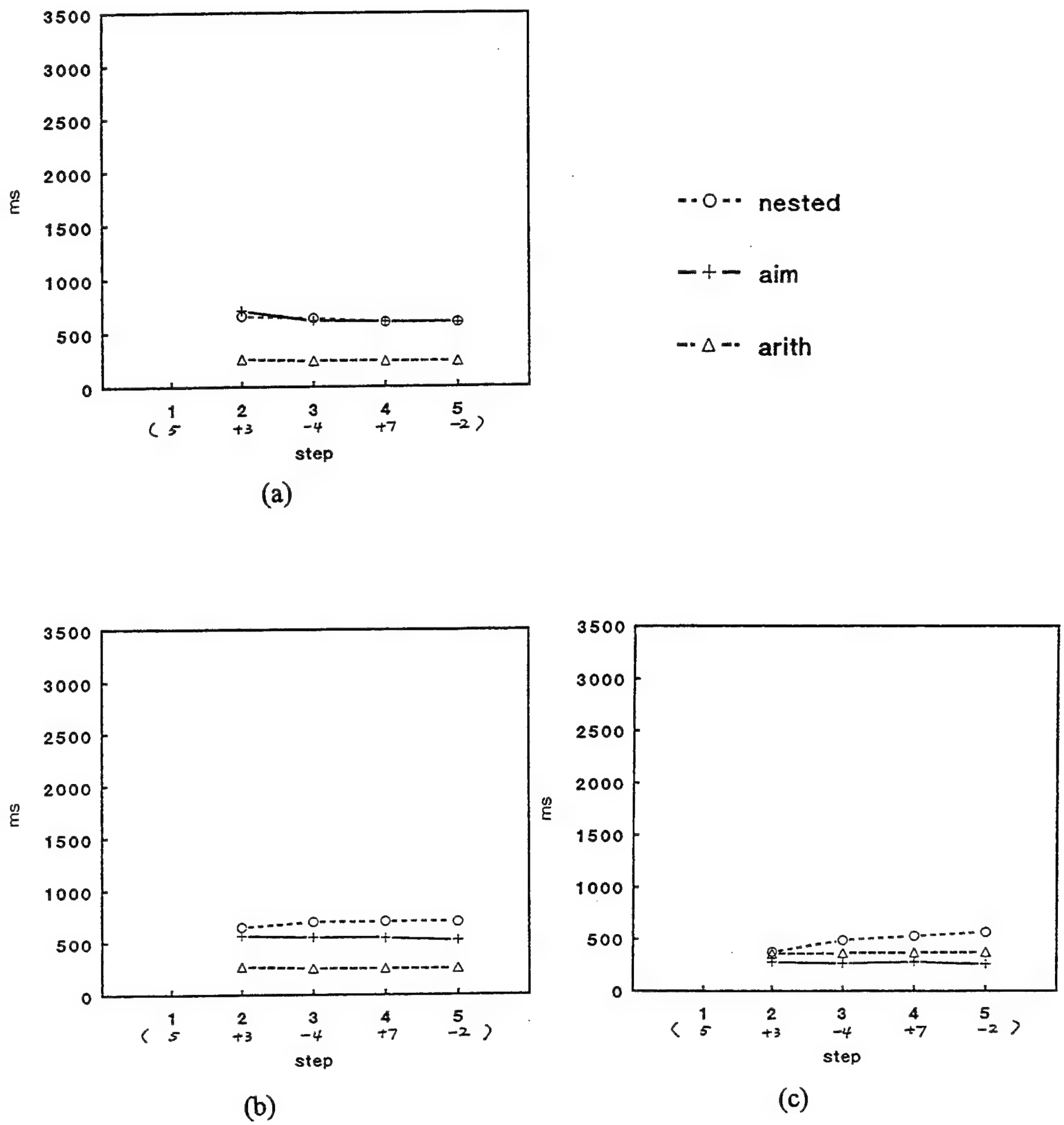


Figure4. Movement time for individual steps as a function of task and step number for (a) the calculation-based movement condition, (b) the visual cue-based movement condition, and (c) no-decision condition for the last quarter of practice.

In the overall condition x task x step analysis of variance on movement time, the main effect of step was not significant, $F < .1$, $p > .5$. However, the effect of step depended on the condition, $F(6,498) = 9.25$, $MSe = 14149$, $p < .0001$. The sequential position effect was found only in the no-decision condition, $F(3,174) = 9.32$, $MSe = 14173$, $p < .0001$, in which movement time was shorter at step 2 than steps 3 through 5, Newman-Keuls, $ps < .01$, which did not differ from one another. In the visual cue-based movement condition, movement time did not differ as a function of step, Newman-Keuls $ps > .4$, and in the calculation-based condition, movement time was *longer* at step 2 than steps 3 through 5, $F(3,150) = 4.88$, $MSe = 26168$, $p < .005$, Newman-Keuls $ps < .05$, which did not differ from one another. That the sequential position effect was absent in the calculation-based movement condition is counter to predictions of the memory-load hypothesis. However, this pattern of results is consistent with the checking hypothesis, because the sequential position effect was found in the phases during which arithmetic processes appear to have taken place.

In summary, the main results of this experiment is as follows. First, in all groups, step time and dwell time was longer in the nested aiming-arithmetic task than in the aiming-alone task and showed similar sequential position effects for all moving conditions including the calculation-based movement conditions. Second, in the calculation-based movement condition, movement time was statistically the same in the nested aiming-arithmetic and aiming-alone tasks; moreover, movement time did not show sequential position effects. Third, in the visual cue-based movement and no-decision conditions, movement time was longer in the nested aiming-arithmetic task than in the aiming-only task, and movement time did show sequential position effects in the no-decision condition.

What do these results reveal about the source of sequential position effects? First, the finding that step time showed similar sequential position effects in all three groups argues against the calculation-postponement hypothesis. Second, it appears that arithmetic calculations took place only during dwell in the calculation-based movement condition but took place during both dwell and movement in the visual cue-based movement and no-decision conditions. That sequential position

effects were found only during the phases of aiming when arithmetic calculations took place argues against the memory-load hypothesis. The overall pattern of results is consistent with the checking hypothesis.

CONCLUSION

The results of this study suggest that the sequential position effect found in the step times, dwell times, and movement times in the nested aiming-arithmetic task of earlier studies reflected checking processes rather than calculation-postponement strategies or memory load per se. These checking processes may involve evaluating and reconstructing completed calculations. Perhaps, such active working-memory management processes are guided by an action plan (Rosenbaum, 1985, 1987; Rosenbaum, Kenny, & Derr, 1983) that represents the structure of the arithmetic procedures necessary to complete the entire arithmetic sequence. Further research will be directed toward understanding the influence of working-memory management processes on aiming and the construction and use of action plans.

REFERENCES

Chaiken, S. R. (1994). The inspection time not studied: Processing speed ability unrelated to psychometric intelligence. *Intelligence, 19*, 295-316.

Embretson, S. E. (1995). The role of working-memory capacity and general control processes in intelligence. Intelligence, 20, 169-189.

Kyllonen, P. C. & Christal, R. E. (1990). Reasoning is (little more than) working-memory capacity? Intelligence, 14, 389-433.

Rosenbaum, D. A. (1985). Motor programming: A review and rescheduling theory. In H. Heuer, U. Kleinbeck, & K.-M. Schmidt (Eds.), *Motor behavior: Programming, control, and acquisition* (pp. 1-33). Berlin: Springer-Verlag.

Rosenbaum, D.A. (1987). Successive approximations to a model of human motor programming. *The Psychology of Learning and Motivation, 21*, 153-182.

Rosenbaum, D. A., Kenny, S., & Derr, M. A. (1983). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 86-102.

Rosenbaum, D. A., & Patashnik, O. (1980). A mental clock setting process revealed by reaction times. In G. E. Stelmach and J. Requin (eds.) Tutorials of Motor Behavior. Amsterdam: North-Holland.

DOES NITRIC OXIDE MEDIATE CIRCULATORY FAILURE INDUCED BY ENVIRONMENTAL HEATING?

Emily B. Skitek
Graduate Student
Department of Pharmacology and Toxicology

University of Kansas
School of Pharmacy
2056 Malott Hall
Lawrence, KS 66045

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC
and
Armstrong Laboratory

August 1996

DOES NITRIC OXIDE MEDIATE CIRCULATORY FAILURE INDUCED BY ENVIRONMENTAL HEATING?

Emily B. Skitek
Graduate Student
Department of Pharmacology and Toxicology
University of Kansas

Abstract

The purpose of this study was to determine whether nitric oxide (NO) contributes to the hypotensive state induced by environmental heating. This was accomplished by competitively inhibiting NO synthesis both before and after environmental heating with a synthetic analog of L-arginine, N^{ω}-nitro-L-arginine methyl ester (L-NAME). Ketamine-anesthetized rats were instrumented for the measurement of arterial blood pressure, ECG, and temperature at five sites. In *Protocol 1*, animals were given L-NAME (10 or 100 mg/kg) or saline i.v., monitored for a period of 10 minutes, and then heated at an ambient temperature of $40 \pm 1°$ C until MAP decreased to 75 mmHg. Heating was then stopped, and the rat was monitored until death. In *Protocol 2*, animals were heated at an ambient temperature of $40 \pm 1°$ C until MAP decreased to 75 mmHg. Heating was then stopped and L-NAME (10 or 100 mg/kg) or saline was immediately administered i.v. There was no difference in survival times between L-NAME and saline-treated rats in either protocol. These results indicate that bolus administration of L-NAME does not reverse hypotension induced by environmental heating, suggesting that excess levels of NO do not mediate this form of circulatory failure.

DOES NITRIC OXIDE MEDIATE CIRCULATORY FAILURE
INDUCED BY ENVIRONMENTAL HEATING?

Emily B. Skitek

Introduction

In mammals, a primary heat loss mechanism during thermal stress is cutaneous vasodilation. This method of thermoregulation routes warm blood to the surface of the body, thus augmenting heat loss to the environment. In mild to moderate thermal stress, arterial blood pressure is maintained despite this cutaneous vasodilation by a concurrent increase in cardiac output and vasoconstriction of the mesenteric vascular beds. However, during extreme thermal stress the body is not able to maintain arterial blood pressure, resulting in heat stroke and subsequent circulatory failure.

The circulatory failure observed during heat stroke is produced, at least in part, by vasodilation of the formerly vasoconstricted mesenteric vascular beds. Concurrent vasodilation of both the cutaneous and mesenteric vasculature produces hypotension by decreasing total peripheral resistance and venous return; the latter also causes a decrease in cardiac output. The physiological mechanism(s) mediating the mesenteric vasodilation and subsequent hypotension which occurs during environmental heat stroke is unknown. However, Gisolfi and his colleagues have demonstrated that the loss of mesenteric vascular resistance is not due to decreases in either sympathetic nerve activity or circulating catecholamines [7], nor to a direct inhibitory effect on the vascular contractile machinery [11].

Recently, it has been suggested that elevated levels of nitric oxide (NO) in the splanchnic circulation might contribute to the mesenteric vasodilation preceding heat stroke. This suggestion is based on the work of Hall and his colleagues, which showed that the level of •NO-heme complex in the portal venous blood of rats increases dramatically during heat stroke [9]. These investigators therefore hypothesized that an enhanced release of NO might contribute to the splanchnic vasodilation seen during heat stroke [9]. Nitric oxide has already been linked to many other forms of shock. Sustained increases in NO synthesis have been shown to mediate the hypotensive states seen endotoxic shock, septic shock, hemorrhagic shock, traumatic shock and anaphylaxic shock [16, 17,12]. From these observations, it seems plausible that NO might also contribute to the hypotension preceding circulatory shock.

Nitric oxide is an endogenous vasodilator molecule that plays an important role in circulatory regulation. It is released basally upon reaction of L-arginine with the enzyme nitric oxide synthase (NOS). This reaction has been found

to occur in vascular endothelial cells [13], vascular smooth muscle cells [2], macrophages [15, 8], and in both central and peripheral neurons [4]. Under physiological conditions, the synthesis and release of NO in vascular endothelial cells acts to modulate organ blood flow and arterial blood pressure [12]. The production of NO can be competitively inhibited by specific analogues of L-arginine, such as N^{ω}-nitro-L-arginine methyl ester (L-NAME) and N-monomethyl-L-arginine (L-NMMA) [12]. Due to the vasodilating property of NO, NO synthesis inhibition causes hypertension in normothermic animals [6, 10] which can be reversed by subsequent addition of excess L-arginine [1].

The purpose of these experiments was to determine whether an increase in nitric oxide production during thermal stress contributes to the hypotension which accompanies heat stroke. To answer this question, a set of experiments was designed in which NO synthesis was blocked by L-NAME administration both before and after rats were subjected to heat stroke.

Methodology

All experiments and animal care were approved by the Institutional Animal Care and Use Committee of Armstrong Laboratory and were conducted according to the National Institutes of Health "Guide for the Care and Use of Laboratory Animals."

Forty-eight male Sprague-Dawley rats (Charles Rivers Laboratories) weighing between 350 and 400 g were used in this study. Prior to experimentation, animals were housed in polycarbonate cages with free access to Purina rodent chow and water. The rats were maintained on a 12h/12h light/dark cycle (lights on at 0600h) in a climate controlled room (ambient temperature = 24.0 ± 0.5°C).

Rats were anesthetized with Ketamine-HCl (Vetalar; 150 mg/kg i.m.), with supplemental doses provided during experimentation. Ketamine administration at this dose level has been shown to produce prolonged surgical anesthesia in Sprague-Dawley rats [5], and previous experiments indicated that ketamine is an appropriate anesthetic for studying the possible contribution of NO to cardiovascular responses induced by thermal stress [30]. A Teflon catheter was placed into a carotid artery for measurement of arterial blood pressure using a Century (Model CP-01) blood pressure transducer connected to a pressure processor (Gould, Model 13-4615-52). A second Teflon catheter was inserted into an external jugular vein for drug infusion. A lead II ECG was obtained by use of nylon-covered fluorocarbon leads attached by a cable to a Gould ECG/Biotach amplifier (Gould, Model 20-4615-65). All measured

variables were recorded continuously throughout experimentation on a Gould TA2000 recorder.

The animals were also instrumented to monitor temperature at five sites: (1) left subcutaneous (lateral, mid-thoracic, side facing away from heat source); (2) right subcutaneous (lateral, mid-thoracic, side facing heat source); (3) left tympanic; (4) colonic (5-6 cm post-anus); and, (5) tail (subcutaneous, dorsal, 1 cm from base). All temperature measurements were obtained via thermistor probes (BSD Medical Corporation) attached to a precision thermometry system (BSD Medical Corporation, Model BSD-200). All temperature and cardiovascular data were A/D (analog-to-digital) converted by an IBM-compatible custom-designed Physiological Monitoring System [3] with real-time graphics display and data analysis capabilities.

Animals were individually exposed to thermal stress in a custom-designed environmental heating chamber in which ambient temperature was monitored and controlled via a thermostat (Tamb). Rats sat on a custon-designed rack consisting of seven 0.5-cm (O.D.) Plexiglas® rods mounted in a semicircular pattern on 4 x 6 cm Plexiglas® plates (0.5 cm thick). Heating was accomplished using commercially available heat sources.

L-NAME was obtained from Sigma Chemical CO. (St. Louis, MO) and solubilized in 0.9% NaCL immediately prior to administration.

Experimental

Protocol 1. This protocol was designed to determine whether pre-treatment with L-NAME, a NOS inhibitor, could attenuate the hypotension induced by subsequent environmental heat stress and thus increase survival time.

After surgery, the anesthetized rat was placed in an environmental heating chamber on a Plexiglas® holder. Rats (n = 8 / group) were monitored at room temperature (25.5 ± 0.3 °C) for a control period of 30 minutes, and then given L-NAME (10 or 100 mg/kg), or saline (vehicle) in an intraveneous bolus injection. After a 10 minute monitoring period, the environmental heating chamber was heated to a temperature of 40 ± 1 ° C and maintained at that temperature until mean arterial pressure (MAP) decreased to 75 mm Hg. Preliminary experiments indicated that if heating is discontinued at this point, MAP will continue to decline until death and the unaided circulatory system can not recover (unpublished observations). Because of this finding, the point at which MAP reached 75 mmHg was defined as the onset of circulatory failure. Therefore heating was stopped at MAP = 75 mmHg, and the rat was monitored until death. Supplemental doses of Ketamine-HCl (Vetalar; 0.1 cc i.m.) were given approximately every 30 minutes throughout the

experiment to maintain sedation. Supplementary Ketamine-HCl injections were postponed if the time of administration fell within 10 minutes of administration of L-NAME.

Protocol 2. This protocol was designed to determine whether post-treatment with L-NAME could attenuate the hypotension induced by environmental heating and thus increase survival time.

After surgery, the anesthetized rat was placed in an environmental heating chamber on a Plexiglas® holder. The rat was monitored at room temperature ($25.4 \pm 0.2\,°C$) for a control period of 30 minutes, the environmental heating chamber was heated to a temperature of $40 \pm 1\,°C$ and maintained at that temperature until mean arterial pressure (MAP) decreased to 75 mm Hg. At this point, heating was stopped and either L-NAME in a low-dose (10mg/kg; $n = 8$), a high dose (100 mg/kg; $n = 8$), or saline (vehicle; $n = 8$) was given intraveneously in a bolus injection. The rat was monitored until death. As described before, supplemental doses of Ketamine-HCl (Vetalar; 0.1 cc i.m.) were given approximately every 30 minutes throughout the experiment to maintain a surgical plane of anesthesia.

Data Analysis

To determine differences between thermal and cardiovascular responses to environmental heating, a two-way analysis of variance (ANOVA) was applied to the three dosage groups, followed by the Student-Newman-Keuls multiple comparison test. A one-way ANOVA was applied to determine differences between the three dosage groups in the heating time required to decrease MAP to 75 mmHg and survival time.

In all of the statistical tests, results were considered significant when $p < 0.05$. All data presented in the results section are means ± standard error.

Results

Protocol 1 (see graphs 1A-1H). During environmental heating, left skin temperature (Tlsk), right skin temperature (Trsk), colonic temperature (Tcol), and left tympanic temperature (Ttym) all rose significantly from control values and continued to increase until heating was stopped. Tail temparature (Tta) increased significantly during initial stages of heating, but this rate of rise slowed after approximately 20 minutes of heat exposure. Thermal responses during environmental heating did not differ between groups with two exceptions; Ttym at MAP = 75 mmHg was significantly different in saline treated rats than in L-NAME treated rats, and Tcol at MAP = 75 mmHg was significantly

different between the saline and low dose groups. Once heating was stopped, Tlsk, Trsk, and Tta began decreasing and continued to decrease until death.

Following the control period, infusion of saline evoked an average increase in MAP of 8.3 ± 5.9 mmHg; L-NAME produced an average initial increase in MAP of 37.6 ± 3.9 mmHg for the low dose (10 mg/kg) and 56.1 ± 5.3 mmHg for the high dose (100 mg/kg). MAP then began declining and continued to decline until death. It should be noted that the average control values for MAP were significantly higher in saline treated rats than rats given L-NAME. However, the average time of heating required to decrease MAP to to 75 mmHg did not differ significantly between groups (see table 1). Also, survival time after cessation of heating did not differ significantly between groups (see table 1). Heart rate increased continuously throughout environmental heating in all groups.

Protocol 2 (see graphs 2A-2H). There was no significant difference in thermal parameters during the control period between groups. During environmental heating, Tlsk, Trsk, Tcol, and Ttym all rose significantly from control values and continued to increase until heating was stopped. Tta increased significantly during initial stages of heating, but this rate of rise slowed after approximately 20 minutes of heat exposure. After heating was stopped, Tlsk, Trsk, and Tta began decreasing and continued to decrease until death.

Cardiovascular parameters during the control period did not differ significantly between dosage groups. During heating, MAP decreased and heart rate increased. The average time of heating required to decrease MAP to to 75 mmHg did not differ significantly between groups (see table 2). Upon cessation of heating, infusion of saline evoked an average increase in MAP of 12.2 ± 4.3 mmHg; L-NAME produced an average initial increase in MAP of 32.3 ± 11.4 mm Hg for the low dose (10 mg/kg) and 24.7 ± 11.0 mm Hg for the high dose (100 mg/kg). There was no significant difference in the average increase in MAP between rats given the low dose or high dose of L-NAME, but the increase in MAP was significantly greater in rats given L-NAME than those given saline. Responses to drug administration were not sustained in any group, and MAP resumed declining rapidly until death. Survival times did not differ between rats given saline and those given L-NAME (see table 2).

Discussion

Based on the results of the present study, it would appear that the elevated levels of NO present during thermal stress do not directly contribute to the development of hypotension during hyperthermia. Although nitric oxide levels

were not measured in these experiments, it has previously been shown in similar rat models of ambient heating that levels of •NO-heme complex consistantly increase at colonic temperatures greater that 39°C [9].

If elevated levels of sphlanchnic NO directly cause the hypotension seen during thermal stress, one would expect to observe an attenuation of this hypotension upon blockade of NO synthesis. Our results show that inhibition of NO synthesis before or after a lethal dose of environmental heating does not effect the cardiovascular response to heat stress or increase survival time. Concurrent with these results, it has previously been demonstrated that hypotension induced by millimeter wave (MMW) heating is not reversed by the subsequent administration of L-NAME [14]. The results obtained in the present and previous study imply that increased levels of NO are not a direct cause of the hypotension seen during thermal stress.

Conclusion

The major finding of this study is that acute pharmacological blockade of NO synthesis either before or after the induction of cardiovascular failure by environmental heating does not attenuate the hypotensive state and does not increase survival time.

References

1. Arnal, J.F., L. Warin, and J.B. Michel. Determinants of aortic cyclic guanosine monophosphate in hypertension induced by chronic inhibition of nitric oxide synthase. *J. Clin. Invest.* 90: 647-52, 1992.

2. Beasley, D., J.H. Schwartz, and B.M. Brenner. Interleukin 1 induces prolonged L-arginine-dependant cyclic guanosine monophosphate and nitrite production in rat vascular smooth muscle cells. *J. Clin. Invest.* 87: 602-8, 1991.

3. Berger, R. E., M. R. Frei, and J. R. Jauchem. Computer-based monitoring of thermoregulatory responses of anesthetized rats exposed to radiofrequency radiation. *Physiologist* 34: 246, 1991.

4. Collier, J., and P. Vallance. Second messenger role for NO widens to nervous and immune systems. *Trends Pharmacol. Sci.* 10: 427-31, 1989.

5. Frei, M. R., and J. R. Jauchem. Effects of 2.8-GHz microwaves on restrained and ketamine-anesthetized rats. *Radiat. Environ. Biophys.* 28: 155-164, 1989.

6. Gardiner, S.M. A.M. Compton, T. Bennett, R.M. Palmer, and S. Moncada. Regional haemodynamic changes during oral ingestion of NG-monomethyl-L-arginine or NG-nitro-L-arginine methyl ester in conscious Bratteboro rats. *Br. J. Pharmacol.* 101: 10-12, 1990.

7. Gisolfi, C. V., R. D. Matthes, K. C. Kregel, and R. Oppliger. Splanchnic sympathetic nerve activity and circulating catecholamines in the hyperthermic rat. *J. Appl. Physiol.* 70: 1821-6, 1991.

8. Granger, D.L., J.B. Hibbs Jr., J.R. Perfect, and D.T. Durack. Metabolic fate of L-arginine in relation to

microbiostatic capability of murine macrophages. *J. Clin. Invest.* 85:264-73, 1990.

9. Hall, D. M., G. R. Buettner, R. D. Matthes, and C. V. Gisolfi. Hyperthermia stimulates nitric oxide formation: electron paramagnetic resonance detection of ·NO-heme in blood. *J. Appl. Physiol.* 77: 548-553, 1994.

10. Konayashi, Y. K. Ikeda, K. Shinozuka, Y. Nara, Y. Yamori, and K. Hattori. L-nitroarginine increases blood pressure in the rat. *Clin.Exp.Pharmacol. Physiol.* 18: 397-9, 1991.

11. Kregel, K. C., and C. V. Gisolfi. Circulatory responses to vasoconstrictor agents during passive heating in the rat. *J. Appl. Physiol.* 68: 1220-7, 1990.

12. Moncada, S., R. M. J. Palmer, and E. A. Higgs. Nitric oxide: Physiology, pathophysiology, and pharmacology. *Pharmacol. Rev.* 43: 109-142, 1991.

13. Palmer, R.M. L-arginine is the physiological precursor for the formation of nitric oxide in endothelium-dependant relaxation. *Biochem. Biophys. Res. Comm.* 153: 1251-6, 1988.

14. Ryan, K. L., M. R. Frei, R. E. Berger, and J. R. Jauchem. Does nitric oxide mediate circulatory failure induced by 35-GHz microwave heating? *Shock*, in press.

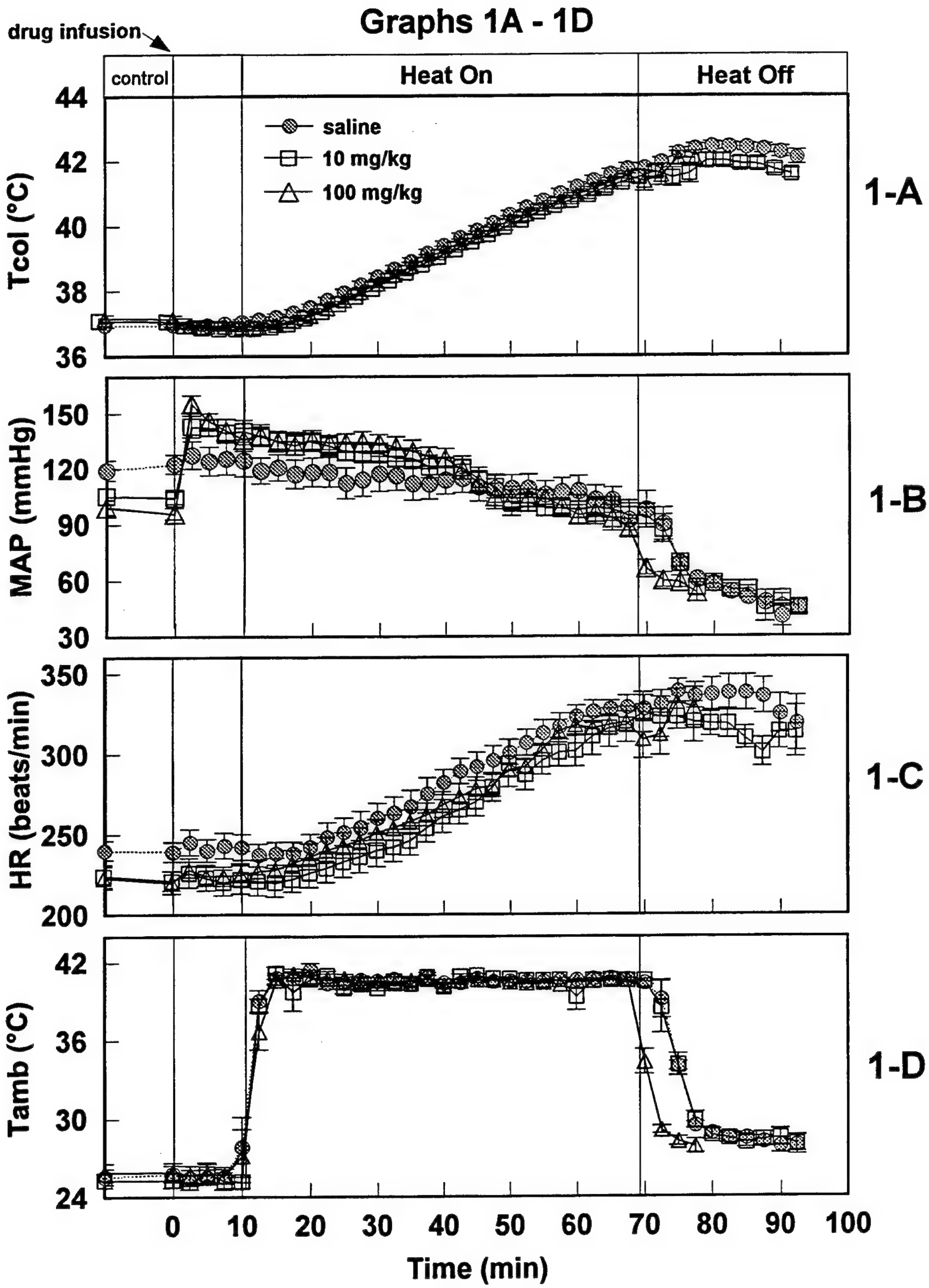
15. Stuehr, D.J., and M.A. Marletta. Mammalian nitrate biosynthesis: mouse macrophages produce nitrite and nitrate in response to Escherichia coli lipopolysaccharide. *Proc. Natl. Acad. Sci. USA.* 82: 7738-42, 1985.

16. Szabó, C. and C. Thiemermann. Invited opinion: Role of nitric oxide in hemorrhagic, traumatic, and anaphylactic shock and thermal injury. *Shock* 2: 145-155, 1994.

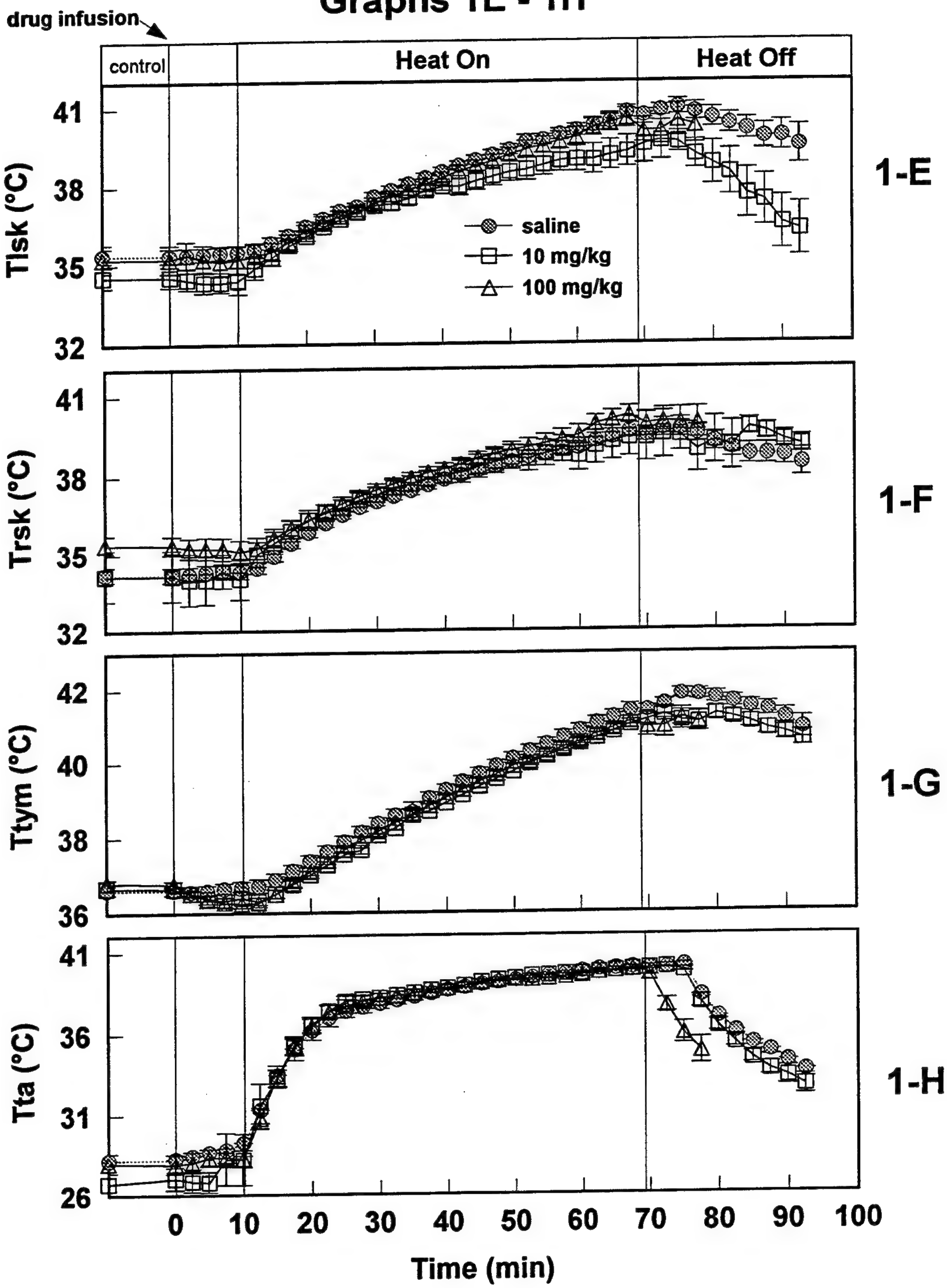
17. Thiemermann, C. The role of the L-arginine: nitric oxide pathway in circulatory shock. *Adv. Pharmacol.* 28: 45-79, 1994.

Table 1: Protocol 1		
	Heating Time (min.) (time for MAP to reach 75 mmHg)	Survival Time (min.)
Saline	64.2 ± 2.1	19.5 ± 3.2
10 mg/kg L-NAME	59.3 ± 4.1	15.5 ± 3.7
100 mg/kg L-NAME	54.1 ± 3.5	9.4 ± 3.4
All Groups	59.2	14.8

Table 2: Protocol 2		
	Heating Time (min.) (time for MAP to reach 75 mmHg)	Survival Time (min.)
Saline	69.4 ± 2.1	14.0 ± 3.8
10 mg/kg L-NAME	64.4 ± 2.8	21.7 ± 7.7
100 mg/kg L-NAME	61.3 ± 4.7	15.0 ± 7.2
All Groups	65.0	16.9

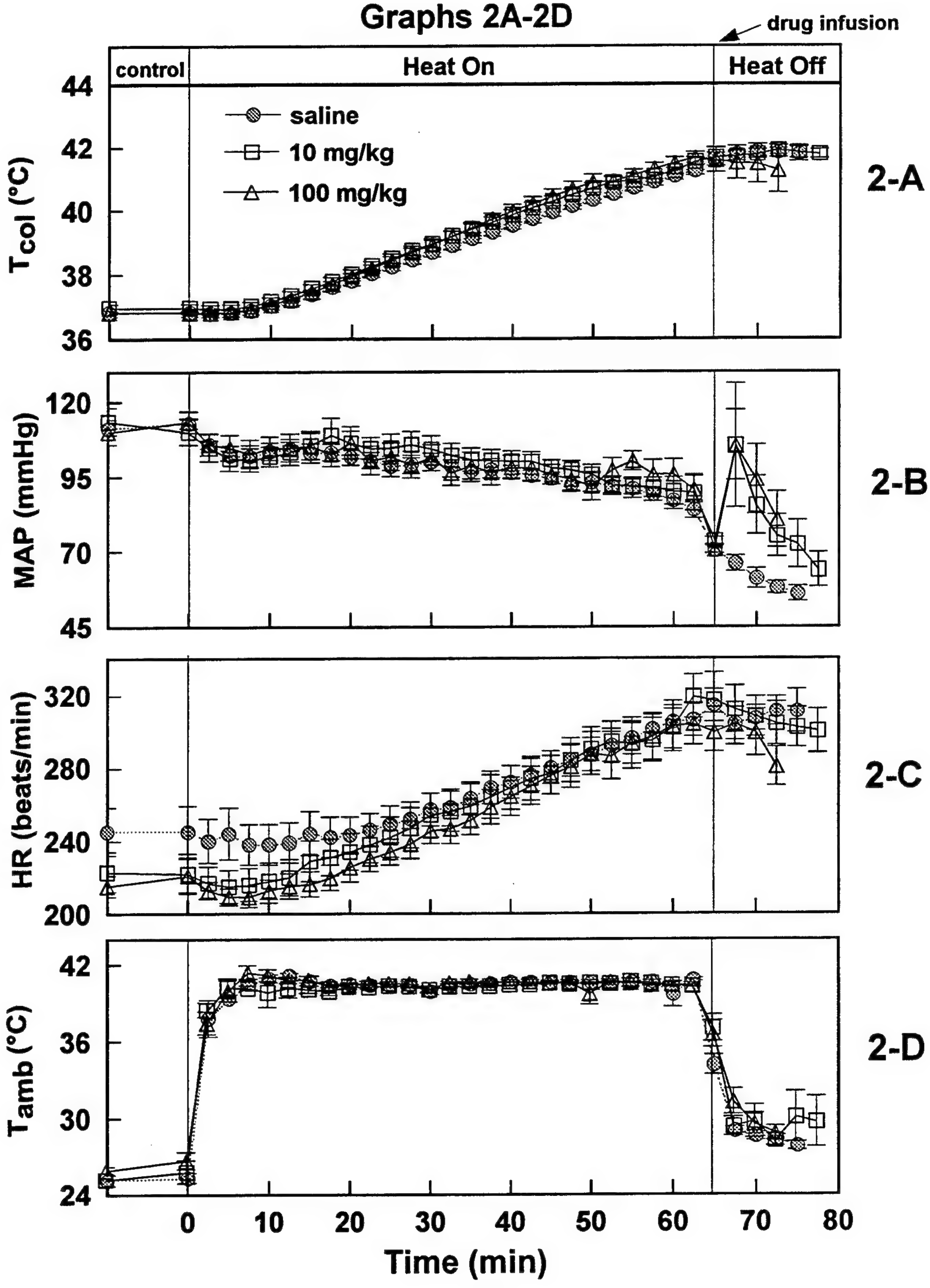

drug infusion
Graphs 1A - 1D
control
Heat On
Heat Off
saline
10 mg/kg
100 mg/kg
Tcol (°C)
44
42
40
38
36
1-A
MAP (mmHg)
150
120
90
60
30
1-B
HR (beats/min)
350
300
250
200
1-C
Tamb (°C)
42
36
30
24
1-D
0
10
20
30
40
50
60
70
80
90
100
Time (min)

Graphs 1E - 1H

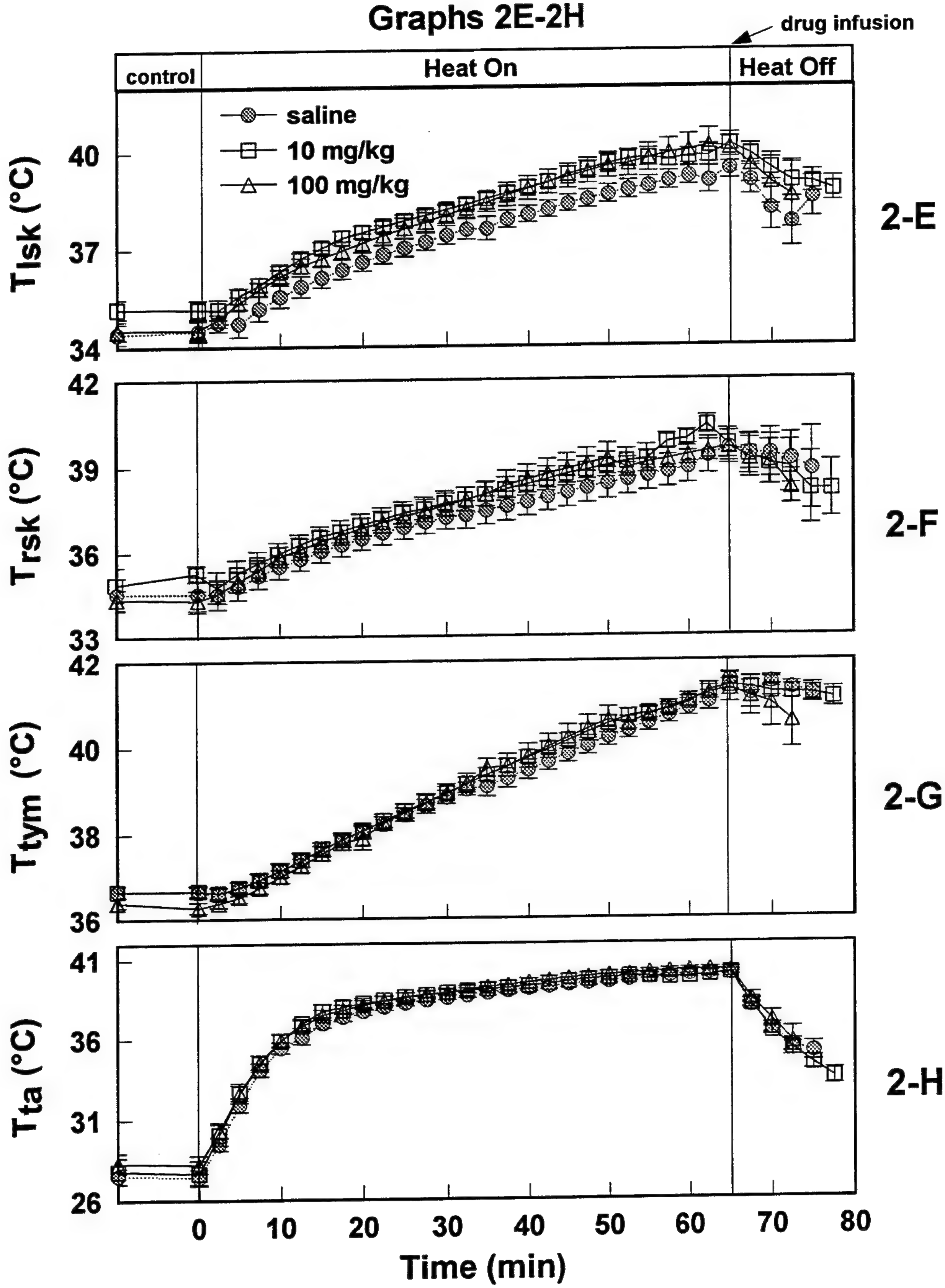


drug infusion
control
Heat On
Heat Off
Tlsk (°C)
41
38
35
32
saline
10 mg/kg
100 mg/kg
1-E
Trsk (°C)
41
38
35
32
1-F
Ttym (°C)
42
40
38
36
1-G
Tta (°C)
41
36
31
26
1-H
0
10
20
30
40
50
60
70
80
90
100
Time (min)

Graphs 2A-2D


drug infusion
control
Heat On
Heat Off
saline
10 mg/kg
100 mg/kg
44
42
40
38
36
T_{col} (°C)
2-A
120
95
70
45
MAP (mmHg)
2-B
320
280
240
200
HR (beats/min)
2-C
42
36
30
24
T_{amb} (°C)
2-D
0
10
20
30
40
50
60
70
80
Time (min)

Graphs 2E-2H


drug infusion
control
Heat On
Heat Off
saline
10 mg/kg
100 mg/kg
Tlsk (°C)
34
37
40
2-E
Trsk (°C)
33
36
39
42
2-F
Ttym (°C)
36
38
40
42
2-G
Tta (°C)
26
31
36
41
2-H
0
10
20
30
40
50
60
70
80
Time (min)

THE DEVELOPMENT OF A GENERAL MEASURE OF PERFORMANCE

Travis C. Tubre
Graduate Student
Department of Psychology

Texas A&M University
College Station, TX 77843-4235

Final Report for
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Brooks Air Force Base, TX

and

Armstrong Laboratory

September 1996

THE DEVELOPMENT OF A GENERAL MEASURE OF PERFORMANCE

Travis C. Tubre
Graduate Student
Department of Psychology
Texas A&M University

Abstract

The U.S. military has invested considerable resources in developing and validating approaches to measuring individual and workgroup performance. However, these approaches have typically been expensive to develop and time consuming to administer. In addition, considerable information about specific job content is often required to develop performance measures using these approaches. This paper describes recent research activities related to the development of a general measure of performance based on recent conceptualizations (e.g., Campbell, 1990a; Borman & Motowidlo, 1993; Viswesvaran, 1993) of the structure of performance which assert that aspects of performance generalize across different jobs. One appealing aspect of such models rests in the ability to develop approaches to measuring and predicting performance which are useful across a broad range of jobs. To date, however, these models have generally been examined at the conceptual level only and have rarely been empirically tested. The present paper describes the development of a core set of items which could be used to (1) empirically test various latent factor models of performance and (2) form the basis for a measure that could be used to obtain general job performance criterion data for a variety of uses (e.g., test validation, program evaluation). Results from the work to date and plans for future activities will be highlighted and discussed.

THE DEVELOPMENT OF A GENERAL MEASURE OF PERFORMANCE

Travis C. Tubre

Introduction

Job performance is, perhaps, the most important construct in industrial and organizational (I/O) psychology and human resource management (HRM). Selection systems attempt to predict performance, while the majority of other organizational interventions (e.g., training, performance appraisal) focus on measuring or improving performance in some way. However, despite its importance, relatively little is known about the latent structure of performance. Indeed, many authors (e.g., Binning & Barrett, 1989; Campbell, 1990a; 1990c) have noted that of the parameters in the classic prediction model, performance has been the most ignored. Any number of predictors (e.g., cognitive ability, conscientiousness) have received a great deal more attention than has been devoted to the criterion. As noted by Viswesvaran (1993), very few efforts have been directed toward developing generalizable models of performance. It has typically been assumed that what constitutes performance differs from job to job. As a result, researchers have used countless numbers of measures as indicators of performance.

More recently, however, researchers (e.g., Borman & Motowidlo, 1993; Campbell, 1990a; Campbell, McCloy, Oppler, & Sager, 1993) have developed theories of job performance which posit that some latent performance dimensions generalize across a broad range of jobs. For instance, Campbell (1990a) asserts that core task proficiency, demonstrating effort, and the maintenance of personal discipline are components of every job. Models that posit the existence of core sets of performance dimensions which exist across a broad range of jobs are appealing for a number of reasons. First, as noted by Campbell (1990c), theory building is becoming an increasingly important component of research in I/O psychology. Since job performance is arguably the most important construct in our domain, a more complete understanding of its structure is a necessity (Viswesvaran, 1993). Second, these theories are consistent with recent models (e.g., Schmidt & Hunter, 1992) describing common causal antecedents of performance (e.g., conscientiousness, general cognitive ability) across jobs of widely varying content. That is,

one plausible explanation for a number of job performance dimensions that generalize across jobs is that a number of variables predict performance in virtually every job. For example, cognitive ability has been shown to predict performance using a variety of criteria across a wide range of jobs. Thus, components of job performance which are strongly related to cognitive ability (e.g., core task proficiency) might be expected to be important components of performance across a broad range of jobs. Finally, if substantiated, such models could provide the basis for developing approaches to measuring and predicting performance which are useful across a variety of jobs. The latter proposition is a primary focus of the present study.

More specifically, the first step in the present study is to empirically test competing conceptualizations of the latent structure of performance. To date, these models have generally been examined at the conceptual level only and have rarely been empirically tested (Campbell, 1990a; Campbell et al., 1993). This paper describes the development of and conceptual basis for an item pool which will be used for this purpose. If performance components which generalize across jobs can be identified, the item pool could also form the basis for a measure that could be used to obtain general job performance criterion data for a variety of uses (e.g., test validation, program evaluation) across a broad range of jobs. Such an instrument could have tremendous utility by reducing the resource demands associated with gathering criterion data.

The Latent Structure of Performance

Viswesvaran (1993) provides an excellent comprehensive review of historical developments in the conceptualization of job performance. As he notes, the literature examining the structure of job performance is fragmented and incomplete. This sentiment has been echoed by scores of researchers (e.g., Arthur & Bennett, 1995; Campbell, 1990a, 1990c) examining aspects of the criterion domain. Early conceptualizations (e.g., Brogden & Taylor, 1950) focused largely on the economic value of individual behaviors to the organization. Measures such as wages, work quality, and supervisory ability were identified (Toops, 1944) and related to the value of an individual to the organization. With the emergence of the literature on expectancy theory, many researchers began to focus on measures that reflected the effort expenditure and productivity of workers (Viswesvaran, 1993). This movement saw the use of criterion measures such as unit production and absenteeism. In the 1970s and 1980s research on prosocial and

organizational citizenship behaviors proliferated (e.g., Bateman & Organ, 1983; Smith, Organ, & Near, 1983). This resulted in the introduction of a variety of criterion measures including teamwork, compliance, and altruism. Finally, in recent years, the impact of counterproductive behavior in the workplace has been studied extensively (e.g., Collins, 1996; Ones, Viswesvaran, & Schmidt, 1993; Sackett, 1994). This literature has yielded a number of criterion measures (e.g., substance abuse, theft, vandalism) related to honesty and integrity in the workplace.

Campbell (1990a; Campbell et al., 1993) provided one of the first large scale attempts to integrate the numerous dimensions of performance into a comprehensive model. According to Campbell, the latent structure of job performance can be modeled using the following eight general factors: (1) job-specific task proficiency, (2) non-job-specific task proficiency, (3) written and oral communication, (4) demonstrating effort, (5) maintaining personal discipline, (6) facilitating peer and team performance, (7) supervision/leadership, and (8) management/administration.

Job-specific task proficiency refers to the degree to which the individual can perform the core substantive or technical tasks that are central to the job and distinguish one job from another. Non-job-specific task proficiency refers to the extent to which the individual performs tasks or executes performance behaviors that are not specific to their particular job but are required of all members of the organization. Written and oral communication is defined as the proficiency with which an individual can communicate (through writing or speech) independent of the correctness of the subject matter. Demonstrating effort reflects the extent to which the individual displays perseverance and intensity in the completion of job tasks. Maintaining personal discipline captures the tendency of the individual to avoid negative behaviors including excessive absenteeism, substance abuse, and rule violations. Facilitating team performance refers to the extent to which the individual performs such behaviors as helping peers, being a good model, and reinforcing the participation of peers. The supervision/leadership dimension focuses on behaviors directed at influencing and maintaining the performance of subordinates. Finally, management/administration includes all of the elements of management that are distinct from direct supervision (e.g., monitoring progress, controlling expenditures).

According to Campbell (1990a; Campbell et al., 1993), these eight factors represent the highest-order factors that can be useful for describing performance in every job in the occupational domain, although some factors may not be relevant for all jobs. As mentioned previously, Campbell contends that core task proficiency, demonstrating effort, and maintaining personal discipline are important components of performance in every job. Campbell's model was largely influenced by the long-term Selection and Classification Project (Project A) sponsored by the U.S. Army (Campbell, 1990b; Campbell & Zook, 1990). While this model represents one of the most comprehensive treatments of the latent structure of job performance currently available, it has rarely been empirically tested. In fact, Campbell et al. (1993, p. 49) admit that direct evidence in support of the model is sparse. In response, they call for future construct validation efforts to test the adequacy of the eight-factor model.

In a similar vein, Borman and Motowidlo (1993) outlined the conceptual and theoretical basis for expanding the criterion domain beyond task (i.e., job-specific) performance to include elements of contextual performance. Drawing heavily from the literature on organizational citizenship behavior (Barnard, 1938; Bateman & Organ, 1983; Smith et al., 1983), prosocial organizational behavior (Brief & Motowidlo, 1986; Graham, 1986; Organ, 1988), and findings from Project A (Campbell, 1990b), Borman and Motowidlo (1993) described the structure of the contextual performance domain (Viswesvaran, 1993). Within this framework, contextual performance is principally defined in terms of behaviors that support the broad organizational, social, and psychological environment of the organization in contrast to behaviors that support the organization's technical core (Borman & Motowidlo, 1993; Motowidlo & Van Scotter, 1994). Contextual performance is further distinguished from task performance in that it is typically more discretionary as opposed to role prescribed. The authors describe five categories of contextual performance as follows: (1) volunteering to carry out task activities that are not formally part of the job, (2) persisting with extra enthusiasm when necessary to successfully complete task activities, (3) helping and cooperating with others, (4) following organizational rules and procedures even when it is personally inconvenient, and (5) endorsing, supporting, and defending organizational objectives.

As with Campbell's (1990a) model of performance, much remains to be accomplished with regard to providing empirical evidence for the adequacy of the task versus contextual performance distinctions. However, the model proposed by Borman and Motowidlo (1993) has recently received empirical support. Motowidlo and Van Scotter (1994) demonstrated that task and contextual performance contributed independently to overall performance in a sample of 421 U.S. Air Force mechanics. Further, their findings suggested that job experience was more highly correlated with task performance than with contextual performance, and personality variables (e.g., dependability) were more predictive of contextual performance than of task performance. These findings are logically consistent with Borman and Motowidlo's (1993) description of task and contextual performance dimensions. That is, within their framework, variation in task performance is posited to reflect individual differences in the proficiency with which task activities are carried out. Thus, individual differences in the knowledge, skills, and abilities associated with a given task should be more predictive of task performance than personality characteristics. Additionally, experience and training performance should be more highly correlated with task performance (Motowidlo & Van Scotter, 1994). Conversely, behaviors such as cooperation, persistence, and compliance would likely be more strongly related to personality variables than to experience, training performance, or ability.

In general, there is substantial overlap between the performance models proposed by Campbell (1990a; Campbell et al., 1993) and Borman and Motowidlo (1993). Campbell's (1990a) job-specific and non-job-specific task proficiency factors are captured in Borman and Motowidlo's (1993) task performance domain. Moreover, the majority of behaviors that Borman and Motowidlo describe as contextual performance would fit into Campbell's (1990a) demonstrating effort, maintaining personal discipline, and facilitating peer and team performance factors. However, it should be noted that Campbell's (1990a) model does not seem to adequately capture the contextual performance dimension associated with endorsing, supporting, and defending organizational objectives.

While somewhat different in their treatment of the criterion domain, neither the model proposed by Campbell (1990a), nor that proposed by Borman and Motowidlo (1993) fully examines the possibility of a general performance factor at the highest level of a hierarchical

structure. In fact, as noted previously, Campbell (1990a) explicitly argues that his eight factors describe the highest order latent variables that can usefully describe performance. In contrast, a model proposed by Viswesvaran (1993) posits the existence of a strong general performance factor which explains substantial variation in virtually all measures of job performance that have appeared in the literature.

Using meta-analytic techniques, Viswesvaran (1993) cumulated studies reporting correlations between various measures of job performance. Next, he grouped the large number of measures into 25 conceptually distinct categories (e.g., quality of performance, communication skills, compliance and acceptance of authority). Based on an extensive literature review, he identified five themes which captured the vast number of performance measures utilized in the literature and sorted the 25 measures into these groups. The groups he utilized are as follows: (1) productivity, (2) conscientiousness, (3) interpersonal skills, (4) withdrawal (e.g., absenteeism, turnover), and (5) measures of overall job performance. Finally, he tested the adequacy of a three-level hierarchical model of job performance with a general performance factor at the highest level, the five-group factors at the second level, and the 25 categories of performance measurements at the lowest level. His results indicated a positive manifold of true score correlations among the 25 performance dimensions (Viswesvaran, 1993). In addition, the three-level hierarchical model provided a better fit to the data than a two-level hierarchical model in which the 25 dimensions were posited to load on a general factor.

The review of the literature on the factor structure of performance provided in this paper indicates that no clear consensus exists concerning the structure of the criterion domain. However, models such as those provided by Campbell (1990a), Borman and Motowidlo (1993), and Viswesvaran (1993) represent a much needed foundation in the development of comprehensive theories of work performance. The next logical step in this process is to empirically test the adequacy of these competing conceptualizations. It is this process which is the central focus of the present study.

Methodology

Testing competing theories concerning the latent structure of job performance necessarily requires the development of an item pool which is representative of the various performance

dimensions presented in the competing theories. The development of this item pool represents the initial step in the present study. Based on an extensive review of such sources as published articles, books, and a variety of instruments designed to measure various dimensions of performance, a 125 item scale was constructed which adequately captures all of the generalizable dimensions of performance specified in the previously mentioned theories. This scale is presented in Appendix A. To ensure accuracy, each stage in the item development process involved the collaborative efforts of two graduate students, a senior faculty member at a large research institution, and a senior research scientist from the U.S. Air Force. Approximately 500 items describing performance in a broad variety of jobs were extracted and modified from sources such as those listed above. Next, each item was examined within the framework of the competing models identified previously. This was done to ensure that every dimension presented in the models was represented by a subset of items. Next, items that were vague or unclear were removed from the pool. Following this, items whose content was extremely similar to other items were removed from the pool. This process was repeated several times to reach consensus among the participants in the process. The final outcome of the item development process is presented in Appendix A.

The next stage of research proposed in the present paper is to empirically test the previously presented theories using the recently developed scale. This process will involve the administration of an adaptive computer-based version of the scale to a large number of U.S. Air Force personnel across a broad range of job categories. Incumbents, in selected occupations, will be asked to rate each of the scale items according to the extent to which each item would be an appropriate measure of performance in their jobs. In addition, data will be collected from supervisors who will rate the extent to which each of the scale items would be an appropriate measure of performance for their subordinates. Representative examples of the instruction sets that will be utilized are also included in Appendix A. It should be noted that the scale does not include items representing job-specific task performance. Rather, for each occupational classification included in the sample, a number of items dealing with job-specific task proficiency will be added to the scale. These items will be drawn from existing performance measurement instruments and job analysis data. In cases where this type of data is nonexistent or outdated, job

analytic techniques will be employed to gather data that will be used for constructing items which measure job-specific task performance.

Following the data collection stage, factor analytic techniques will be utilized to assess the adequacy of various models describing the latent structure of job performance. This analytic process will also be used for the purpose of scale refinement. In addition, the nature of the data set will facilitate the testing of hypotheses concerning the stability of the factor structure of performance across different levels of occupations and incumbent characteristics (e.g., sex). Further, administering versions of the scale which incorporate or omit job-specific content will allow for the testing of various research questions concerning the relative impact associated with the inclusion or exclusion of job-specific content. To increase the generalizability of our findings, we will attempt to incorporate a large number of occupational classes that are analogous to jobs in the civilian sector. In addition, data will be collected from incumbents in jobs with varying degrees of technical content.

It is important to note, however, that the goals of the present study extend beyond simply conducting empirical tests of competing theories of job performance. If our initial analyses indicate that various dimensions of performance are generalizable across a broad range of jobs, the scale items could form the basis for a measure that could be used to obtain general job performance criterion data for a variety of uses (e.g., test validation, program evaluation). Depending on the nature of the criterion data of interest, the scale could be modified to include job-specific content. The U.S. military has invested considerable resources in developing and validating approaches to measuring individual and workgroup performance which are largely based on specific job content. However, these approaches have typically been expensive to develop and time consuming to administer. Thus, a general measure of performance which could be modified to obtain general criterion data across a broad range of jobs would be of tremendous utility. The development of such an instrument is the long term goal of our efforts thus far. However, before moving on to the next stage, questions concerning the latent structure of performance must be addressed.

Conclusions

As many authors (e.g., Arthur & Bennett, 1995; Viswesvaran, 1993) have noted, the existing literature on the criterion domain is fragmented and incomplete. This study represents an attempt to clarify some of the confusion concerning what constitutes successful or unsuccessful job performance. The theories presented by Campbell (1990a), Borman and Motowidlo (1993), and Viswesvaran (1993) are the groundwork upon which the present study will build. The importance of a more complete understanding of the construct of job performance cannot be overstated. Job performance is in some way related to virtually every activity carried out within the context of I/O psychology and HRM. However, our understanding of its structure and composition lags behind our understanding of predictors and outcomes of successful job performance. In addition, efforts directed at identifying generalizable dimensions are particularly valuable. As noted by Viswesvaran (1993):

> Developing theories of job performance for each task (or even job) will hinder the development of a general theoretical understanding of the construct of job performance. As the content generality of the dimensions increases, the value of the dimensions in developing prediction instruments and theories of work-behavior increases (p. 64).

However, it is our belief that the value of developing theories of job performance incorporating dimensions with increased content generality goes beyond the development of prediction instruments and theories of work-behavior. Empirically identifying dimensions of performance with generalizable content represents the first step in the development of instruments that could be used to obtain general criterion data across a broad range of jobs. Instruments of this type could substantially reduce the resource demands of criterion measurement. In addition, such instruments would expand the criterion domain to include elements of contextual performance which may be overlooked in more traditional job-specific approaches to criterion measurement.

References

Arthur, W., Jr., & Bennett, W. R.., Jr., (1995). The international assignee: The relative importance of factors perceived to contribute to success. Personnel Psychology, 48, 99-114.

Barnard, C. (1938). The functions of the executive. Cambridge, MA: Harvard University Press.

Bateman, T. S., & Organ, D. W. (1983). Job satisfaction and the good soldier: The relationship between affect and employee "citizenship". Academy of Management Journal, 26, 587-595.

Binning, J. F., & Barrett, G. V. (1989). Validity of personnel decisions: A conceptual analysis of the inferential and evidential bases. Journal of Applied Psychology, 74, 478-494.

Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), Personnel selection in organizations (pp. 71-98). San Francisco, CA: Jossey Bass.

Brief, A. P., & Motowidlo, S. J. (1986). Prosocial organizational behaviors. Academy of Management Review, 11, 710-725.

Brogden, H. E., & Taylor, E. K. (1950). The dollar criterion: Applying the cost accounting concept to criterion construction. Personnel Psychology, 3, 133-154.

Campbell, J. P. (1990a). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), Handbook of Industrial and Organizational Psychology (2nd Ed., Vol. 1, pp. 687-732). Palo Alto, CA: Consulting Psychologists Press.

Campbell, J. P. (1990b). An overview of the army selection and classification project. Personnel Psychology, 43, 231-239.

Campbell, J. P. (1990c). The role of theory in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), Handbook of Industrial and Organizational Psychology (2nd. Ed., Vol. 1, pp. 39-73). Palo Alto, CA: Consulting Psychologists Press.

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1992). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), Personnel Selection in Organizations (pp. 35-70). San Francisco, CA: Jossey Bass.

Campbell, J. P., & Zook, L. M. (Eds.). (1990). Improving the selection, classification, and utilization of army enlisted personnel: Final report on Project A. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Collins, J. M. (1996). A narrative and empirical evaluation of the socialization trait as a predominant predictor of productive and counterproductive work behavior. Manuscript submitted for publication.

Graham, J. W. (1986, August). Organizational citizenship informed by political theory. Paper presented at the meeting of the Academy of Management, Chicago, IL.

Ones, Viswesvaran, C., & Schmidt, F. L. (1993). Meta-analysis of integrity test validities. Journal of Applied Psychology, 78, 679-703.

Organ, D. W. (1988). Organizational citizenship behavior: The good soldier syndrome. Lexington, MA: Lexington Books.

Motowidlo, S. J., & Van Scotter, J. R. (1994). Evidence that task performance should be distinguished from contextual performance. Journal of Applied Psychology, 79, 475-480.

Sackett, P. R. (1994). Integrity testing for personnel selection. Current Directions in Psychological Science, 3, 73-76.

Schmidt, F. L., & Hunter, J. E. (1992). Development of a causal model of processes determining job performance. Current Directions in Psychological Science, 1, 89-92.

Smith, C. A., Organ, D.W., & Near, J. P. (1983). Organizational citizenship behavior: Its nature and antecedents. Journal of Applied Psychology, 68, 653-663.

Toops, H. A. (1944). The criterion. Educational and Psychological Measurement, 4, 271-297.

Viswesvaran, C. (1993). Modeling job performance: Is there a general factor? Unpublished doctoral dissertation, University of Iowa.

APPENDIX A

General Performance Measurement Item Pool

Instructions (Incumbents)

Please rate the following items on the extent to which they would be applicable to measuring job performance in your occupation. Your responses should not indicate the level at which you are currently performing the behavior associated with each item. Rather, you should indicate the appropriateness of each item as a measure of performance for employees in your occupation.

Instructions (Supervisors)

Please rate the following items on the extent to which they would be applicable to measuring the job performance of your subordinates. Your responses should not indicate the level at which your subordinates are currently performing the behavior associated with each item. Rather, you should indicate the appropriateness of each item as a measure of performance for your subordinates.

1. Accepts direction from supervisors
2. Challenges conventional thinking
3. Contributes innovative ideas
4. Deliberates before making important decisions
5. Delivers effective customer service
6. Determines appropriate courses of action
7. Diagnoses problems
8. Displays commitment to the job
9. Displays flexibility in work assignments
10. Grasps new information/tasks quickly
11. Maintains a positive attitude
12. Makes effective and timely decisions
13. Requires minimal supervision

14. Responds well to feedback
15. Sets realistic goals
16. Shows a clear understanding of the job
17. Understands the needs of clients
18. Uses criticism constructively
19. Works well independently
20. Works well under adverse or uncertain conditions
21. Clearly expresses ideas and concepts in writing
22. Clearly expresses ideas and concepts orally
23. Delivers clear, well organized presentations
24. Displays impressive communication skills
25. Effectively leads discussions
26. Effectively presents complex information
27. Efficiently prepares written materials
28. Gets his/her point across
29. Interacts well with others
30. Presents material in a logical order
31. Speaks clearly and concisely
32. Speaks well in front of a group
33. Uses a convincing presentation style
34. Commits personal time to the job if necessary
35. Displays determination when faced with difficult tasks
36. Follows through consistently
37. Gets work done quickly
38. Is highly motivated to accomplish tasks
39. Maintains a high level of output
40. Meets agreed upon deadlines
41. Pursues all job functions with energy
42. Readily accepts new responsibilities

43. Responds immediately when required
44. Responds to challenging circumstances
45. Shows a strong desire to excel
46. Strives to meet goals of the unit
47. Takes on additional responsibilities in emergencies
48. Volunteers for extra duties
49. Works effectively under stressful conditions
50. Works extra hours when necessary
51. Avoids excessive absenteeism
52. Avoids substance abuse
53. Complies with company policies and procedures
54. Exhibits honesty in the work environment
55. Has high standards of personal integrity
56. Maintains stable relationships
57. Meets financial obligations
58. Obeys rules and regulations
59. Preserves confidentiality
60. Reports instances of unethical behavior
61. Reports to work on time
62. Shows discipline in the work place
63. Treats others with respect
64. Acknowledges the contributions of peers
65. Addresses and resolves conflicts
66. Assists co-workers in task completion
67. Clarifies ambiguous roles
68. Cooperates well with others
69. Demonstrates interpersonal sensitivity
70. Encourages participation among co-workers
71. Garners the support of co-workers

72. Identifies with the needs of co-workers
73. Inspires peers to excel
74. Provides assistance to team members
75. Provides constructive feedback
76. Provides encouragement/guidance to co-workers
77. Provides training to others
78. Readily accepts the opinions/suggestions of others
79. Resolves conflicts
80. Seeks out the opinions of others
81. Shares expertise and experience with others
82. Shows consideration for the feelings of fellow workers
83. Treats others fairly and consistently
84. Challenges workers to excel
85. Clearly conveys expectations for assignments
86. Conveys a sense of purpose and mission
87. Demonstrates authority and knowledge
88. Develops the skills of subordinates
89. Displays prudence in supervision
90. Exhibits a strong sense of capability
91. Generates a feeling of energy
92. Generates a sense of personal involvement
93. Helps others to achieve their full potential
94. Identifies strengths/development needs in others
95. Inspires others to take action
96. Maintains discipline in the work place
97. Motivates subordinates to accomplish goals
98. Plays a leadership role
99. Promotes a positive image of the organization
100. Provides subordinates with a sense of job-security

101. Provides useful and accurate feedback
102. Takes appropriate risks to achieve results
103. Takes decisive action when warranted
104. Uses effective leadership skills
105. Accomplishes goals through cooperation
106. Addresses conflict in the unit
107. Attracts and selects high caliber talent
108. Clarifies roles and responsibilities
109. Differentiates between good and poor job performance
110. Effectively delegates responsibility
111. Effectively manages organization-wide activities
112. Exhibits effective administrative skills
113. Facilitates the discussion and resolution of different views
114. Integrates diverse efforts
115. Manages resources effectively
116. Monitors efforts and takes corrective action
117. Observes and evaluates other employees
118. Provides useful feedback
119. Represents the unit well
120. Translates strategies into objectives
121. Understands causes of problems
122. Understands organizational priorities
123. Presents a positive image of the organization to others
124. Defends organizational objectives
125. Supports the directives of the unit

Renardo D. Tyner report not available at time of publication.

THE ANALYSIS OF AQUEOUS FILM FORMING FOAM

Christopher G. Walker
Jackson State University
Jackson, MS 39216

Final Report
Summer Graduate Research Program
Armstrong Laboratory

Sponsored By
The Office of Scientific Research

August, 1996

THE ANALYSIS OF AQUEOUS FILM FORMING FOAMS

Christopher G. Walker
Jackson State University

Abstract

Analytical techniques were employed to determine the specific organic components in aqueous film forming foams (AFFF). Electrophoresis was used extensively in an attempt to identify and separate surfactants and non surfactant organic components. In this study we investigated a ion pairing solid phase extraction technique (SPE) to cleanup and separate methylene blue active components. The complexes are then concentrated and derivatized with diazomethane followed by GC/MS analysis. Further laboratory work is needed to definitively characterize the organic constituents in AFFF. The results and limitations of these techniques are discussed.

THE ANALYSIS OF AQUEOUS
FILM FORMING FOAMS

Background

The Air Force is current interested in treating wastewater generated from fire fighting training bases. These wastewater contain residual fuel and aqueous film foam (AFFF) constituents. Several adverse affects upon both sewage treatment systems and aquatic environments.[1] This has lead to a reduction in fire training activities at several Air Force and Navy facilities. The Air Force is currently interested in treating the wastewater generated from fire fighting training exercises.

The research problem at Air Forces Bases involve the potential of on site systems as an effective method for treatment of these wastewaters.[2] Several studies have examined the degradation of jet fuel, however until recently only few studies have examined the components of AFFF. Most of these studies used nonspecific methods. The determination of anionic surfactants in effluents is relied heavily on specific complexing of the anionic species with a fluorescent dye, rhodamine, 6G, followed by fluorescence detection or by reaction with methylene blue and calorimetric UV-Vis spectrophotometric detection. These techniques can yield artificially high values because they lack the specificity required to monitor fluorosurfactants in the presence of other anionic surfactants and UV absorbing chemicals.[4] Methods such as COD, BOD and TOC does not measure AFFF directly. Since residual jet fuel can contribute significantly to the measurements, the percentage of biodegraded AFFF is usually underestimated.[5] Additionally, HPLC procedures have been complicated due to the ionic character of these compounds. These procedures lack either specificity, sufficient analyte resolution, or sensitivity.

Objective of the Research

The overall objective of the Aqueous Film Forming Foam project is to development a method using existing on site instrumentation to gain a better understand of the constituents in AFFF. Although attempts have been made to analyze the fate of AFFF's, there are no known analytical procedures for measuring AFFF components. Afterwards, use these methods to evaluate the effectiveness of on -site biological treatment to remove AFFF containing wastewater.

The specific goals of my project were to:

1) Investigate electrophoresis as a potential technique for the analysis of AFFF's and it's constituents.

2) Identify methods that could possibly be used to separate and quantify the organic components of AFFF's

AFFF Composition

Patent #3,773,195 issued November, 1973 to Minnesota Mining and Manufacturing Company spells out, in very general terms, the composition of AFFFs. In general, the composition of AFFFs may be comprised of:

1. A water-soluble fluoroaliphatic surfactant (s).

2. A water-soluble synthetic fluorine-free surfactant (s) which is hydrocarbon congruous.

3. A water stabilizer selected from the group consisting of alkyl ethers of alkylene glycol and glycerol such as polyethylene glycol (Polyox N-10).

4. Water

The fluorinated surfactant has the formula:

$$(Rf)_n \; Qm \; Z$$

whereas:

n	=	1 or 2
Rf	=	a fluorinated saturated monovalent organic radical having a terminal perfluoromethyl group containing 3 to 20 carbon atoms.
Qm	=	a multivalent linking group and may contain divalent oxygen or trivalent nitrogen atoms, bonded only to carbon atoms.
m	=	an integer from 0 to 2.
z	=	a water solublizing polar group.

Capillary Electrophoresis

Introduction

Ansul Company Ansulite and 3M FC-203/206 are the AFFF's mostly being used by the Air Force. [5] After initial review of the information available on AFFF's, we began laboratory work using FC-203 as a model AFFF because of information about its components and possible method of analysis was

available from 3M. The components of interest in AFFF are the fluoroalkyl surfactants. These components are responsible for roughly 10% of total volume of AFFF. (Table 1) Several attempts have been made to analyze of AFFF via chromatographic techniques with little success.

Table 1. COMPONENTS OF 3M AFFF-FC-203

Ingredients	%v/v
Butyl Carbytol	15
Water	76
Fluoroalkyl Surfactants (2)	5
Hydrocarbon Surfactants (2)	5
Urea	4
Tolytriazole	<0.1

Capillary electrophoresis and related separation modes have become important techniques for the separation of peptides, proteins, and related compounds for several reasons. The key reasons are that; it is applicable to a small sample volume and complex and multicomponent samples, and it is capable of developing extraordinary high efficiency because of plug flow and elimination of mass-transfer factors since a second phase is absent. As a result little reagent is consumed. Another advantage is that CE can also be used to separate classes of smaller sized analytes of those with higher mobilities, including inorganic anions and cations. Furthermore, the scope of CE applications is broadened by deactivation of the silanol sites on the silica capillary wall. For example, using a low pH buffer represses silanol dissociation and shields.

Capillary zone electrophoresis has proven to be a very effective analytical tool and is ideally suited for the separation of ionized compounds like organic and inorganic ions. Separations in capillary are based on differences in the mobilities of charged solutes in response to an applied electric field. The electrophoretic mobility on an ion is largely determined by its charge, mass, and molecular geometry. The

migration velocity of an ion is also influenced by electroosmotic flow (EOF), which refers to the bulk movement of electrolyte within the capillary.

While CE offers a viable alternative for the separation of anionic species, certain limitations must be considered when designing a separation strategy. The high electroosmotic flow inherent with a capillary columns hindered the analysis of complex anions in solution by creating bulk fluid movement toward the injector side of the capillary column when a negative potential is applied. A positive potential must be avoided if the goal of the experiment is to attract the anionic species past the detector window, since the anode is on the injector side of the column and the electrophoretic mobility of the anions will be toward the anode. The best way to solve this equivocation of polarity dependent mobility and EOF is by suppressing the EOF. Other works have described these EOF suppression or reversal by addition of modifiers such as quaternary ammonium salts and alcohols. In the present work, several approaches are taken. CIA Pak OFM Anion-BT was employed to suppress the EOF while using a capillary column. The separation of anionic components is based purely on electrophoretic mobility of the analytes since EOF is greatly suppressed. Unwanted neutral species and cations migrate to injector side of the column because on minimal EOF and reverse polarity. Since the anions migrate toward the detector side of the column, a bi directional system exists. This bi-directional systems provides a simplified process for separating and detecting the analytes of interest over traditional techniques such as HPLC and GC.

Some fluorosurfactants, components of AFFF, are UV in-active. That is these components have little or no UV absorbance. For this reason an indirect UV detection scheme was employed for the analysis of fluorosurfactants. Indirect UV detection is simple and can be performed with ease. Indirect UV detection uses a UV-absorbing ionic chromophore in the electrolyte which is displaced by the analyte ion resulting in a drop in absorbance. Several buffers were used to identify a carrier electrolyte that would achieve the optimum separation.

Experimental

The CE system used for this study was a Beckman P/ACE 5500 Capillary Electrophoresis instrument, equipped with UV-Detector and System Gold software integration package (Beckman

Instruments, Fullertion, CA). A System/CAP capillary column with a 57 cm length and 75μm inner diameter from Beckman. The detector window was 50 cm from the injection end of the capillary column. The CE electrical configuration was in reverse and normal polarity mode for anion detection using a constant current mode ranging from 10-30μA. Indirect and direct UV detection was utilized during the analysis. In each run method the capillary columns were conditioned by rinsing with a regeneration solution (1% Sodium Hydroxide) for 4 minutes, followed by a 5 minute rinse with DI water, and a 10 minute rinse with the run buffer solution. A 5 minute rinse of run buffer solution preceded and followed each sample run.

Reagents

All standard and samples solutions were prepared with Milli Q grade water and filtered with #1 ashless filter paper. CIA-Pak OFM Anion-BT buffer solution was obtained from Water's Technical Division (Millford, MA). All other reagents were obtained from Fisher Scientific (Saint Louis, MO).

Results and Discussions

Several run method were developed to maximize the retention time and peak efficiency for each electrolyte. The electropherograms in Figure 1 were obtain after several weeks of method development (varying run time, current, and temperature.) The data was collected at 200, 254, 280, and 214nm for AFFF (Ansulite and 3M), FC-99, FC-95 (surfactants 3M), and butyl carbytol. Chromatographic analysis was perform in attempt to gain qualitative information on the AFFF's its composition. After several weeks of analysis with CE and UV-Visible we tentatively identified butyl carbytol as peak I at (2.75 minutes) 200nm in (Figure 2). This was also confirmed by GC/MS. The chromatographic profile of AFFF-*3M* and AFFF-*Ansulite* in Figure 3 had some similarities, however these similarities are probably due to urea and neutral organic artifacts since the retention time is near that of the electroosmotic flow marker. Several pH's were experimented to investigate the effects of pH on the chromatographic profile of AFFFs and surfactants . These experiments did not reveal any significant qualitative information.

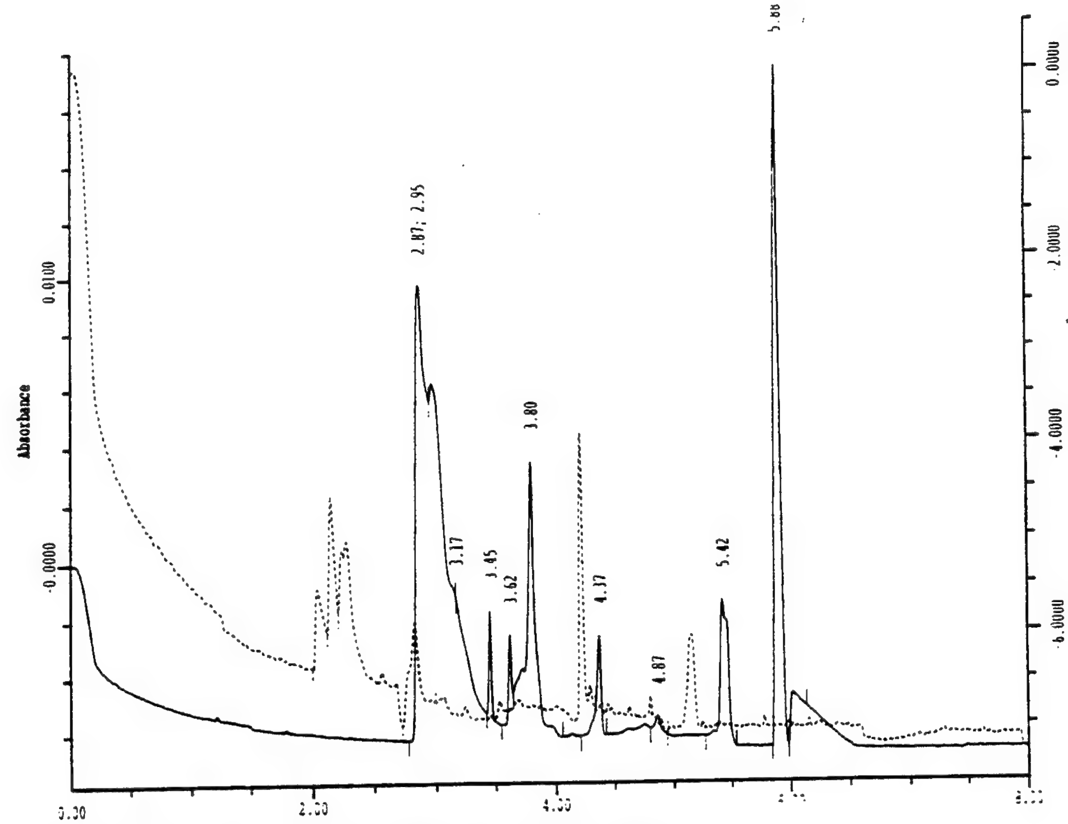


Figure 1. FC-99,3M surfactant (----) and AFFF-Ansulite at 200nm.

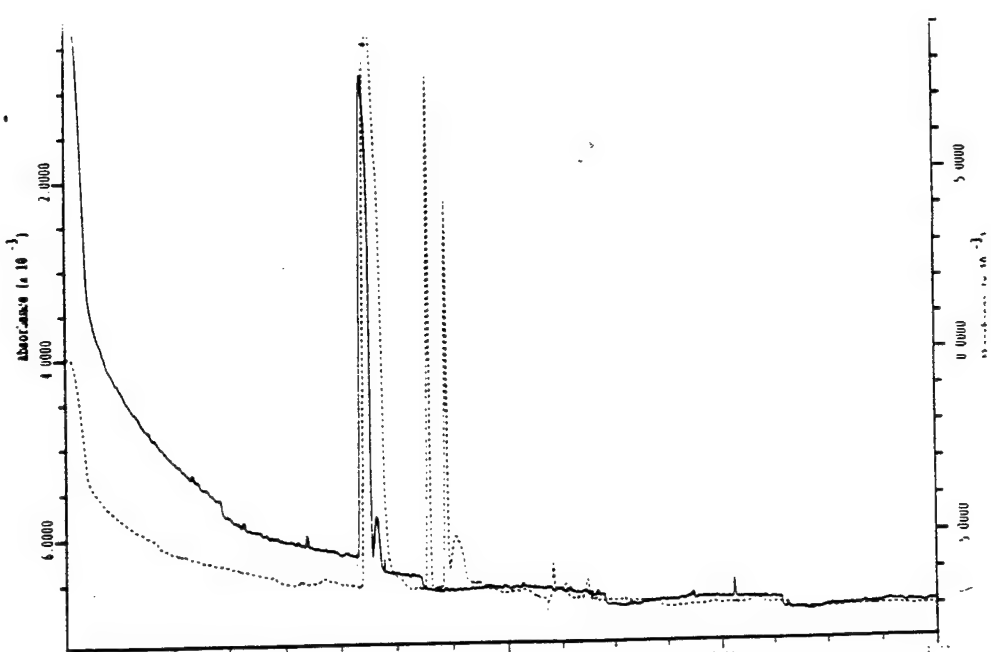


Figure 2. Butyl Carbytol (-----) & AFFF-3M (—). Detected at 200nm.

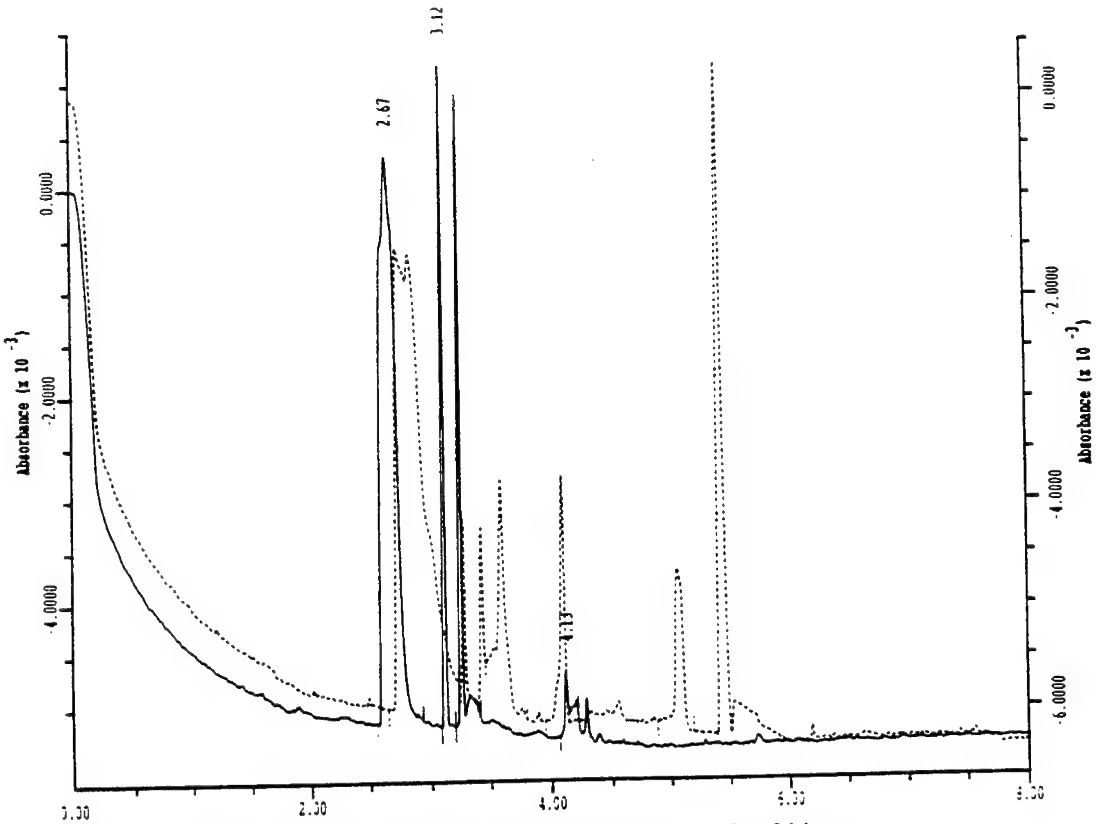


Figure 3. AFFF-3M (----) and AFFF-Ansulite (—). Detected at 214nm.

After analyzing UV-Visible spectrometry and CE data of the fluorosurfactants it was concluded that the peaks in the electropherograms were not a result of the fluorosurfactants. The fluorosurfactants have little or no UV absorbance at the wavelengths within the UV range of the detector. As a result of the lack of UV absorbance of the fluorosurfactants, we decided to employ indirect UV detection 254 with a chromate buffer. This method had superior resolution, however we were unable to characterize any component. After a conference with Water's technical service representatives we decided to attempt to separate the anionic species by the suppression electroosmotic flow marker. Water's OFM BT solution was prepared with a sodium tetraborate buffer in an attempt to eliminated neutrals which were present and bias to UV detection. Standards were prepared with neutral, cationic, and anionic species. The neutral and cationic UV interfering chemicals were eliminated from the electrophoretic system by reverse polarity and minimal EOF allow for the detection of the anionic species. Any EOF generated by the reverse polarity mode was in the direction of the injector, causing the neutral and cationic species to exit the

capillary without passing the detector window. Even though this method allow the detection of anionic species, we were unable to analyze AFFF and related standards. This was thought to be due to the large molecular weight of AFFF and the C9 - C18 standards. Several other electrolytes were use in the analysis with marginal results.

Ion Pair Solid Phase Extraction and Derivatization GC/MS

Introduction

A wide variety of methods have been used to determine surfactants in both commercial products and environmental samples. The majority of analytical method for LAS determination in waste water and water employed ion pair formation of LAS with methylene blue and spectrophotometric determination. Methylene blue was used as an ion pairing reagent in the analysis of anionic surfactants. Methylene blue, a cationic dye, pairs with anion substances in aqueous solutions forming an immiscible organic solid upon equilibration. The MBAS method is useful for estimating the anionic surfactant content of waters and wastewater, however it is not fully specific for anionic surfactants. Other organic and inorganic anions also form ion with methylene blue. For the specific characterization of individual LAS gas chromatography mass spectrometry must be applied. Ion pair formation with methylene blue does allow the routine extraction of the LAS components from complex mixtures of organic compounds. In this study we investigated a ion pairing solid phase extraction technique (SPE) to cleanup and separate methylene blue active components of the extracts for GC/MS analysis (figure 2). The advantage of GC/MS is that complex mixtures can be resolved into individual components. Anionic surfactants are recovered by methylene blue complexation on and C-18 column and eluted with chloroform. SPE has been widely used for isolating aliphatic compounds from aqueous environmental samples. The complexes are then concentrated and derivatized with diazomethane for GC/MS analysis.

Experimental Section

GC Chromatography-Mass Spectrometry (GC/MS) was performed on Hewlett-Packard 5970B GC/MS system equipped with a 30m x .25mm I.D. X 0.25μ DB-5 MS capillary column. Samples were

prepared by a derivatization procedure with 50 drops of diazomethame. The GC capillary injection port was at 210° C.

Ion pair formation with methylene blue.

↓

Transfer blue aqueous solution onto SPE column
C-18 column

↓

Elute MBAS with chloroform.

↓

Evaporate to dryness

↓

Derivatize with diazomethane

↓

Concentrate, and analyze by GC/MS.

Figure 4. Flow diagram of ion pair SPE and derivatization procedure.

The temperature was held at 40° C for 5 minutes, rammed 10° a minute, to 240° and held at 240° for 30 minutes. The mass spectrometer was scanned . Data was acquired on a HP Chemstation personal computer.

Conclusion

A variety of solvents containing TBA were evaluated for recovering sulfonic acids from the C18 SPE disks. Large volumes of organic solvents were required to elute the surfactants from the C18 disk. Chloroform was determined to be the better solvent for the elution of the surfactants. GC/MS data did not indicate derivatization as a result of the derivatizating agent diazomethane. As a result of the excess ion paring reagent present the quantity of diazomethane was thought to be insufficient to react with the sulfonated surfactants.

RECOMMENDATIONS

1. The research should be continued to identify a column to optimize the separation of the fluorosurfactants.

2. Investigate a buffer system that will be compatible with the column without exaggerating the electroosmotic flow.

3. Investigate other detection methods.

4. Develop a sample preparation procedure whereas the stabilizing agent is eliminated.

5. Identify a derivatizing agent that is effective in the presence of ion-pairing reagents during GC/MS analysis.

Bibliography

1. Fink, P. T., Air Force Technical Report. *Aqueous Film Forming Foams Treatability*, Civil and Environmental Engineering Development Office, Tyndall Air Force Base, FL, 1978.

2. Ventullo, R.M., *Analysis and Fate of Organic Components of Aqueous Film Forming Foams*, Final Report, USAF-UES Summer Faculty Research Program, 1986.

3. Chain, E.S., Wu T.P., and Rowland, R.W.., *Membrane Treatment of Aqueous Film Forming Foam*, Summer Faculty, Georgia Institute of Technology, 1980.

4. Setzkron, E.A., and Carel, A.B., *JAOCS,* Volume 40, 1963.

5. Otsuki, A., and Shiraishi, H., *Analytical Chemistry.,* Volume 51, 1979. p. 2329.

AUTOMATING THE COGNITIVE TASK ANALYSIS PROCEDURE

Ross E. Willis
Graduate Student
Department of Psychology

Texas Tech University
P.O. Box 2051
Lubbock, TX 79409

Final Report for:
Graduate Student Research Program
Armstrong Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington DC
and Armstrong Laboratory

August 1996

AUTOMATING THE COGNITIVE TASK ANALYSIS PROCEDURE

Ross E. Willis
Graduate Student
Department of Psychology
Texas Tech University

Abstract

Research has shown that students learning from private human tutors learn more efficiently than students learning from a traditional classroom setting (Bloom, 1984). However, private human tutors are not always a viable option. It has been proposed that computer-based tutorials can be used to instruct material as effectively as private human tutors. In an effort to accomplish this, computer-based tutorials must evolve toward intelligent tutoring systems, in which the tutoring system has expertise in the domain being instructed. Subject matter expertise may come from analyzing documents within the domain. However, the primary source of knowledge is obtained through interviews and observations of experts. The time consuming process of eliciting knowledge from experts and organizing the knowledge into comprehensive curricula is referred to as cognitive task analysis (CTA). Shute and Torreano (1995) propose an automated cognitive task analysis procedure designed to elicit comprehensive curricula which encompasses expert knowledge and skills in a more timely manner, compared to existing CTA methodologies. The current paper reviews background literature in the domain of CTA and presents a prototype of an automated CTA procedure.

Automating the Cognitive Task Analysis Procedure

Ross E. Willis

Introduction

Shute and Torreano (1995) outline some ideas for an automated cognitive task analysis procedure called DNA (i.e., Decompose, Network, and Assess). The goals of the proposed task analysis methodology are to create an easy-to-use cognitive task analysis procedure which is capable of eliciting comprehensive curricula from experts across many domains. During the course of my summer internship at Armstrong Laboratory (AL/HRTI), Dr. Shute, Dr. Sugrue, and I refined the ideas and methodology associated with the automated cognitive task analysis procedure outlined in the Shute and Torreano paper, as well as developed a prototype of an automated cognitive task analysis procedure. Before actual development of a prototype system, we conducted an extensive literature review and discussed many theoretical issues. The current paper describes the literature of intelligent tutoring systems, existing cognitive task analysis methodologies, and the prototype automated cognitive task analysis procedure.

Brief History of Intelligent Tutoring Systems

It has been noted that students learn more effectively when learning from private tutors compared to learning in standard classrooms (Bloom, 1984). Bloom, for example, reported students learning from private tutors performed an average of 2 standard deviations (i.e., 98^{th} percentile) above students who learned from standard classrooms. It is important to identify characteristics of effective human tutors to produce superior student learning outcomes.

Gamble and Page (1980) noted that effective human tutors: (1) converge student heuristics to those of the tutor; (2) choose appropriate examples and problems for the student; (3) work novel examples; (4) adjust to accommodate different students' backgrounds (e.g., aptitudes, learning styles, personalities); (5) measure progress; and (6) review previously learned material.

The primary disadvantage of private tutors, however, is that they are often expensive and it is not feasible to provide private tutors to all students (Anderson, 1990). Regian and Shute (1992) posit that artificial intelligence techniques can potentially make it possible to develop intelligent tutoring systems which can emulate effective human

tutors and achieve superior learning outcomes similar to those reported by Bloom (1984).

The field of computer-assisted instruction (CAI) is approaching the "intelligence" required to emulate an effective human tutor. Burns and Capps (1988) list a three-stage test of intelligence for CAI to be considered "intelligent". First, the system must "know" the subject matter (i.e., what the system proposes to teach students). In essence, the system must model the knowledge of an expert in the domain. This component of the instructional system is often referred to as the "expert model". Second, the system must model the student's prior knowledge and skills, as well as knowledge and skills that emerge during the course of instruction (i.e., what the student knew prior to instruction and what the student has learned). This component of the instructional system is referred to as the "student model". Third, the system must contain an intelligent strategy that can reduce the difference between the expert model and the student model. This component of the system is called the "tutor". The tutor determines a student's path of advancement throughout the instructional curriculum by comparing the student model to the expert model and making diagnostic decisions based on differences (VanLehn, 1988).

If an automated instructional system (i.e., CAI) adheres to the conditions of this three-part intelligence test, the system is referred to as an intelligent tutoring system (ITS). Although each module of an ITS (i.e., expert model, student model, and tutor) is an important component to the efficiency and effectiveness of the tutorial, this paper primarily focuses on the expert model. Specifically, the question, "How is expert knowledge of the domain obtained?" will be addressed.

Task Analysis

Instructional designers typically obtain knowledge structures for ITSs from subject matter experts (SMEs). The process of acquiring expert knowledge from a SME is termed "task analysis". Traditional task analysis primarily focuses on observable behavioral criteria, such as procedures. Thus, this form of task analysis is often referred to as behavioral task analysis. Behavioral task analysis is limited because knowledge structures are ignored, which makes it difficult to identify cognitive components that novices (i.e., students) need in order to progress towards expertise (Redding, 1989).

Cognitive task analysis (CTA) focuses on underlying knowledge structures that are required for expertise in a given domain (Redding,

1989). CTA improves upon behavioral task analysis in several ways. First, CTA aids in the understanding of how skills and knowledge are obtained and how the learning process can be improved. Second, CTA typically analyzes the knowledge and skills needed to perform the entire task and examines interrelations among these knowledge and skill components. For example, a novice must understand that "Σ" means "the sum of" before a novice can use the formula "$(\Sigma X)/N$" to calculate the mean. Thus, the underlying connection (interrelation) between "Σ" and the formula "$(\Sigma X)/N$" is made apparent. Third, CTA emphasizes skill development. That is, attention is paid to how novices become experts. Thus, knowledge is obtained from experts as well as novices. Fourth, CTA posits that it is often difficult for experts to articulate knowledge and skill components because they may have automated some tasks. That is, experts who perform certain procedures relatively often compile their knowledge of these procedures, which makes it difficult to recall the specific steps of the procedure and reasoning for each step. For example, an expert typist may not be able to verbally recall exactly where the keys are located, but an expert's typing efficiency demonstrates that he or she clearly knows where the keys are located (Anderson, 1990).

The primary objective of cognitive task analysis is to acquire a rich body of knowledge about a domain from experts and compile that knowledge into a useful curriculum to serve as the expert model for the ITS (Durkin, 1995). Anderson (1988) referred to the expert model as the "backbone of any ITS" (p. 21). That is, an ITS cannot effectively instruct a domain without a firm grasp of the domain knowledge. Unfortunately, it is often the case where the expert models ITSs are incomplete and are limited to instructing only a portion of the domain knowledge (Anderson, 1988). Without a rich knowledge base, an ITS will not give efficient instruction. Thus, it is important to develop methods which are capable of effectively capturing a complete array of expert knowledge.

Types of Knowledge

Expert knowledge is characterized by both declarative knowledge and procedural skill (Shute, 1995). Experts are typically able to explicate knowledge about components in a system, whether the system is physical (e.g., automobile), or cognitive (e.g., statistics), how these components operate, and the underlying connections between components. Thus, Shute takes the typical declarative knowledge/procedural skill

framework and decomposes it into three, more descriptive types of knowledge distinctions.

The most fundamental form of knowledge is termed "symbolic knowledge" (SK). SK includes definitions of terms. For example, knowing that the symbol "Σ" means "the sum of".

Procedural skill (PS) refers to the application of some procedure. For example, the ability to calculate the arithmetic mean of a set of data using the formula "$(\Sigma X)/N$". Thus, PS typically requires some prerequisite SK (e.g., one must possess the knowledge of the symbol "Σ" before the formula "$(\Sigma X)/N$" can be calculated).

Conceptual knowledge (CK) refers to organized sets of concepts, rules, and their linkages. CK is the ability to construct higher-level relationships among concepts. For example, knowing that the formula to calculate the mean refers to a measure of central tendency which should be used to summarize quantitative and normal data in a distribution (as opposed to the mode or median). Thus, CK is much more complex than SK or PS, and in many cases, requires prerequisite knowledge of SK and PS.

Sources of Expert Knowledge

Where does the curriculum for an ITS originate? The first phase of the cognitive task analysis procedure is to acquire knowledge about the subject matter; a process known as "knowledge acquisition". One method of obtaining domain-specific knowledge is by reviewing documents such as books, reports, and technical manuals. This method is referred to as document analysis. However, the primary source of expert knowledge for most ITSs is via interviews and observations of the subject matter expert (Durkin, 1994).

The process of eliciting knowledge from experts, on the surface, may seem like a straightforward task; however, appearances can be deceiving. Hayes-Roth, Waterman, and Lenat (1983) refer to the cognitive task analysis procedure as a "bottleneck" in development of an ITS. Since the person conducting the CTA (typically referred to as the knowledge engineer - KE) has much less domain-specific knowledge than the expert, communicating relevant issues becomes challenging. A second challenge lies within the fact that experts typically find it difficult to verbalize much of their knowledge (Gordon, Schmierer, & Gill, 1993; Durkin, 1994; Hall, Gott, & Pokorny, 1995; Moore & Gordon, 1988).

Document analysis. The process of document analysis usually consists of reviewing training manuals, books, technical manuals, and other domain-related literature to determine the required knowledge and

skills. Gordon, Babbitt, Sorensen, Bell, and Crane (1993) mention three problems with document analysis. First, documents may be written ambiguously. That is, documents are often unclear and not concise. Second, some knowledge types are not represented. Documents usually focus on terms and definitions (i.e., SK) and procedures (i.e., PS), and often omit underlying rationale associated with procedures (i.e., CK) (Gordon, Schmierer, & Gill, 1993). Third, documents do not make the relationships between terms apparent.

Document analysis provides a good starting-point because the KE can learn a great deal about a domain in a relatively short period of time. However, document analysis does not provide enough information about a domain to create a well-rounded curriculum. Therefore, it is necessary to obtain information directly from experts.

Interviews. The interviewing method typically involves the KE asking the SME questions related to the domain. The questions are specifically worded to elicit knowledge about the domain. Two types of interviews exist, which differ in terms of the specificity of the questions which are asked.

The unstructured interview is typically used when the KE desires to obtain general knowledge about a domain. Questions in an unstructured interview are usually open-ended and allow the expert to discuss a topic in a natural manner. For example, "What do you know about ______?". An advantage of the unstructured interview is that the knowledge obtained provides a general overview of the domain. However, unstructured interviews often provide information that is fragmented and superficial (i.e., some knowledge and skill components are typically omitted, which makes it difficult to create a rich curriculum).

In contrast to the unstructured interview, which is intended to provide a general overview of the domain, the structured interview technique is designed to provide details about specific issues, in sequence. Questions are posed which elicit information about specific details. Before progressing to other details, the KE collects all data about the knowledge and skills associated with the current detail. This form of questioning is referred to as the funnel sequence (McGraw & Harbison-Briggs, 1989). The funnel sequence begins by asking general questions that address the domain in a broad manner and progresses toward more specific questions.

For example, if a KE was questioning a Blackjack expert, the first question might be, "What do you know that makes you good at Blackjack?". This broad question is designed to define a high-level goal and might

elicit the more specific SME response that he or she understands a card counting strategy. In an effort to obtain more specific information about this sub-part of the domain "Blackjack", the KE poses a more specific question designed to decompose the various card counting strategies.

One advantage of the structured interview is that important issues are defined and decomposed into their constituent sub-parts. A great amount of detail is extracted from the expert when employing the structured interview. A disadvantage is that much attention is paid to guiding the expert to decompose sub-parts that have been explicated. This method often leaves some higher-level goals undefined. For example, a Blackjack expert may expend a great deal of energy decomposing the knowledge about card counting strategies and forget to include information about when to make a "double-down" wager.

The primary disadvantage with structured and unstructured interviewing techniques is that experts find it difficult to verbalize their knowledge about the domain (Durkin, 1994; Redding, 1989; Hall, Gott, & Pokorny, 1995; Gordon, Schmierer, & Gill, 1993; Collins, 1985). Knowledge and skills that experts use, but are unaware of is referred to as implicit knowledge. Collins (1985) reported that experts may not report parts of their knowledge about a domain because they are unaware that they have the knowledge, not because they cannot verbalize it.

Observation. Observation techniques involve acquiring knowledge by having an expert solve problems or actively participate in a domain. The theory is that underlying principles driving decisions will emerge by having experts perform a task in context; thus converting implicit knowledge to explicit knowledge. That is, KEs attempt to induce conceptual knowledge by examining behavior during observation. Two methods of observing expert behavior exist: retrospective and concurrent.

The retrospective observation technique involves having experts solve a problem, reflect on their problem-solving methods, and identify the knowledge and skills used. Retrospective observation is advantageous because it does not interfere with problem solving activity, which has been noted as creating an unnatural problem solving situation (Evans, 1988). That is, an expert solving a problem is not interrupted by a KE asking questions about procedures being used. A disadvantage of retrospective observation is that it requires memory recall which may lead to inaccurate of incomplete reports.

The concurrent observational technique involves having an expert think-aloud while performing a given task. The KE actively participates by encouraging an expert to verbalize the procedures he or she is performing. A benefit of this method is that it does not require an expert to reflect on procedures used during the task. That is, the procedures are explicated at the time they are being used.

Two variations of retrospective and concurrent observational techniques require an expert to solve familiar problems or novel problems. Familiar problems are those which are common to a field. An advantage of selecting familiar problems is that it can provide insight into problem-solving methods employed by experts in the domain. However, since a case is typical, experts may have encountered the problem many times prior to the knowledge acquisition process and may have automated the problem-solving procedure, thus rendering it difficult to recall the actual procedures. That is, experts will often overlook details they may consider irrelevant or unimportant, which will leave gaps in knowledge and skills about the domain.

Novel problems combat the disadvantage of familiar problems (i.e., problem-solving procedures become automated). Durkin (1994) notes that experts, when faced with a novel situation, will resort to more basic problem-solving behaviors, resulting in more complete knowledge structures. Durkin mention that novel problems aide in explicating implicit knowledge. A disadvantage of selecting novel problems, however, is in the actual selection of the novel situation. Often the KE does not have enough knowledge of the domain to generate challenging and novel situations.

Organizing Knowledge and Skills

As previously mentioned, document analysis, interviews, and observations, collectively, are commonly referred to as the "knowledge acquisition" phase of CTA. Once the knowledge acquisition phase is completed, the KE must transcribe and organize all data that were collected. Durkin (1994) mentions that transcription and organization can take as long as 40 hours of effort for every one hour spent collecting data in the knowledge acquisition phase. Generally, KEs video- or audio-tape record interview and observation sessions with an expert. After the session, the recordings are analyzed and transcribed into several pages of notes and graphs of the knowledge and skills elicited from experts. Converting these pages of text and graphs to an

ITS curriculum is an arduous task. Learning hierarchies and conceptual graphs have been used to generate curricula (Gagné & Medsker, 1996).

Learning hierarchies. Learning hierarchies (see Figure 1) are graphical representations of all tasks and corresponding prerequisite subtasks (Gagné & Medsker, 1996). This graphical representation illustrates the interrelations of superordinate goals and prerequisite subordinate goals.

Figure 1 is an incomplete learning hierarchy for the domain of Blackjack devised by conducting a CTA with Dr. Shute serving as a Blackjack expert and Dr. Sugrue and myself serving as KEs. Learning hierarchies contain boxes which are referred to as "nodes". Each node contains a brief label related to a knowledge or skill. Nodes are joined by lines which indicate interrelations among knowledge and skill components. As an illustration of the hierarchy, notice that the superordinate goal of "rules of the game" (e.g., the object of Blackjack is to come closer to 21 than the dealer without exceeding this value) is comprised of subordinate goals such as knowing "definitions of cards and values" (e.g., face cards have values of 10). Students must know "definitions of cards and values" before they can understand "rules of the game".

Because learning hierarchies are generated by decomposing superordinate goals into subordinate goals, they generally contain more complex goals (i.e., CK) near the top of the graph and basic units of knowledge (i.e., SK) toward the bottom of the graph. When designing curricula, instructors would present more basic forms of knowledge such as definitions of cards and values, which is a SK unit, before presenting more complex forms of knowledge such as counting cards, which would be more complex and labeled a CK unit. Thus, curricula construction begins from the bottom of the learning hierarchy and progress upwards. The progression of teaching from SK to CK is based on learning theory (e.g., Anderson, 1987; Kyllonen & Christal, 1989; Kyllonen & Shute, 1989) as well as empirical evidence (Shute, 1995).

Conceptual graphs. Conceptual graphs (see Figure 2) illustrate internal knowledge structures by using nodes that are linked together to show interrelations (Moore & Gordon, 1988).

Figure 2 is an incomplete conceptual graph relevant to searching a videotape for specific information adapted from Gordon, Schmierer, and Gill (1993). Notice that each node contains information about knowledge category. Gordon, Schmierer, and Gill define five node content categories: states, events, styles, goals, goals/actions. States are

definitions of terms (e.g., "fast forward button is labeled 'FF'"). States correspond to SK (i.e., definitions of terms); however, it appears that Gordon, Schmierer, and Gill also use "concept" to refer to SK elements. Events are procedures (e.g., "VCR goes into fast forward mode") and correspond to PS. Styles denote the specific manner in which an event occurs (e.g., "slowly"). Goals are states which are desired (e.g., "VCR turned off"). Goals/Actions specify the action which one must take to accomplish a specified goal (e.g., "search the videotape"). Goals/Actions correspond to CK.

Links among the nodes are directional arcs which indicate prerequisite knowledge for superordinate goals. For example, in order to understand the goal "search forward", a student must first understand the action "press fast forward button", as well as where the fast forward button is located and how the button is labeled.

Existing CTA Methodologies

Several CTA methodologies currently exist (see Hall, Gott, & Pokorny, 1995 for a more complete review). Two popular approaches include PARI (Hall, Gott, & Pokorny, 1990) and conceptual graph analysis (Moore & Gordon, 1988).

PARI. PARI (i.e., Precursors, Actions, Results, and Interpretations)is a structured interviewing technique in which KEs present problems for experts to solve. During the solution of the problem, experts are probed to think-aloud, which provides information about the procedures used to obtain the solution as well as the complex reasoning underlying the problem solving decisions. This methodology is much like the concurrent observation previously described. Thus, the PARI procedure yields information about the precursors to the action taken by experts, as well as the results of those actions and experts' interpretations of those results.

Conceptual graph analysis. Conceptual graph analysis (CGA) is a technique which involves analyzing documents within a given domain and constructing preliminary conceptual graphs of the domain. The next step of CGA involves interviewing experts. Interviews typically involve asking structured "Why, How, What" questions. Examples of a "Why, How, What" questions includes, "Why did you ____?", "How do you _____?", and "What is a _____?". The "Why, How, What" questioning methodology was originally developed by Graesser & Clark (1985), and can be used to elicit SK (i.e., "What" questions), PS (i.e., "How" questions), and CK (i.e., "Why" questions) knowledge structures. In addition to structured

questioning, KEs also present the preliminary conceptual graphs to experts and have them add new information, delete incorrect information, and clarify existing information. The third step is to collect observation data which is used to infer underlying principles driving experts' decisions during problem solving.

The Challenge of CTA

PARI and CGA are only two of many CGA methodologies which exist. Each methodology presents its own strengths and weaknesses. However, the challenge of CTA lies within developing a methodology that is: (a) domain-independent, (b) effectively captures all types of knowledge, (c)shows relationships among knowledge and skill components, (d) is easy to use, and (e) efficient.

A CTA procedure that is domain-independent has the ability to be employed to analyze many domains. Some existing CTA methods are applicable to only certain types of tasks or specific domains. For example, the PARI methodology is almost exclusively used for analyzing technical tasks, such as electrical troubleshooting. The conceptual graph analysis methodology has been reported to be extremely time consuming when used to analyze relatively complex domains (Gordon, Schmierer, & Gill, 1993).

A CTA procedure that effectively captures all types of knowledge (i.e., SK, PS, CK), is capable of providing comprehensive curricula for ITSs. Collecting knowledge and skills from each type of knowledge provides definitions of components within the system, how those components operate, how components operate in relation to other components within the system, and why components operate the way they do.

A CTA procedure that can illustrate the relationship among knowledge components (i.e., hierarchically and conceptually), has the advantage of being able to map easily into curricula. That is, generating a curriculum from learning hierarchies and conceptual graphs is easier than compiling transcriptions from interviews with experts (Gordon, Schmierer, & Gill, 1993).

Most current CTA methodologies are difficult to use. For example, PARI and CGA require training in statistics and cognitive science in order to effectively interpret the data that are collected. Redding (1989) mentions that most people who design curricula (e.g., teachers) do not have training in these areas, thus will not be able to use these methods.

Most current CTA methodologies require a great deal of time to complete. For example, Gordon, Babbitt, Sorensen, Bell, and Crane (1993) report that the CGA required a total of 360 hours to analyze how F-16 pilots use radar during air intercept. Specifically, Gordon et al. (1993) reported 80 hours spent conducting and organizing information obtained during document analysis, 160 hours conducting and organizing information obtained during interviews with nine F-16 expert pilots, and 120 hours collecting and organizing observation data.

Toward an Automated CTA Procedure

Researchers have proposed that automating the CTA procedure could provide a solution to the bottleneck that CTA presents to ITS development (Shute & Torreano, 1995; Redding, 1988). The proposed CTA design, ideas, and methodology was originally outlined by Shute and Torreano (1995). DNA is designed to elicit knowledge from a SME which will decompose (i.e., the "D" in DNA) a domain, network (i.e., the "N" in DNA) the knowledge into comprehensive structures, and employ other experts in a given domain to assess (i.e., the "A" in DNA) the validity and comprehensiveness of the knowledge structures. Goals of the proposed task analysis methodology are to create an easy-to-use cognitive task analysis procedure which is capable of eliciting comprehensive curricula from experts across many domains. The prototype of DNA is a series of interactive modules that automate knowledge acquisition and are based on theoretical and empirical support previously reviewed.

The first module of DNA (i.e., Customize module) is completed by the instructional designer (ID). The ID provides information such as the domain to be decomposed by SMEs and characteristics of the learning population. This information provides the point at which the SME should start decomposing (i.e., superordinate goal) and the point at which the SME should stop decomposing (i.e., the lowest-level subordinate goal). After providing the required information, the Customize module generates a personalized letter explaining the purpose of the project to prospective SMEs and a floppy diskette which will be mailed to prospective SMEs. The diskette contains files for a SME to install on his or her computer which DNA needs to elicit and store knowledge structures.

Once a SME receives and installs the files from DNA diskette, he or she begins the second module (i.e., Explicit Decompose module), which conducts a structured, interactive dialog with an SME specifically

designed to elicit most of the explicit knowledge associated with a domain. However, provisions are included which attempt to elicit any implicit knowledge an expert may recall. DNA utilizes the "Why, How, What" questioning procedure which has been shown to successfully elicit knowledge from experts (Gordon, Schmierer, & Gill, 1993). Questions are designed to elicit SK, PS, and CK. For example, Symbolic knowledge is obtained via questions such as. "What is a definition of ______?". Conceptual knowledge is obtained via questions such as, "Why is _______ important?". Procedural skills are obtained via questions similar to, "How do you _____?", or "What is the first step you do when you ______?".

In addition to the "Why, How, What" questioning technique, the program utilizes a structured interviewing technique in a modified funnel sequence. That is, questions are designed to elicit general superordinate goals first, then specific knowledge related to these superordinate goals. However, our system modifies the funnel sequence by allowing experts the choice of decomposing knowledge in depth or breadth manner. That is, experts are not forced to generate all superordinate goals before decomposing into subordinate goals. Experts will be allowed to generate a superordinate goal and decompose it at any point during the knowledge acquisition phase. Thus, the funnel sequence is not strictly enforced in order to allow customization of use. Questions are iterated until the expert has exhausted domain knowledge.

In an effort to acquire implicit knowledge (i.e., PS that has become automatic), experts will be prompted to list procedures they utilize, answer questions as to why they perform certain steps, and under which conditions these steps are performed. Experts will also be asked to submit with their data, a few common situations, as well as a few novel situations, they have encountered during their careers. These situations will be presented to other experts and solved in a subsequent module.

All information given by experts will be stored in a database as a list of curriculum elements (CEs) which are components of knowledge needed in order to develop expertise in the domain. By storing information in a CE record, it becomes easy to translate the information into teachable units and generate curricula.

The third module (i.e., Network module) loads each CE from the CE record generated in the Explicit Decompose module and allows experts to arrange and link them. That is, experts will generate learning hierarchies and conceptual graphs. Each node contains the name of the CE

and its contents as defined during the Explicit Decompose module. Linkages differ in terms of strength (i.e., weak, moderate, strong), which indicates the degree to which items are related, and directionality (i.e., uni-, or bi-directional), which indicates which CEs are prerequisites for other CEs. The graphical representation is designed to make missing knowledge components apparent to experts, as well as IDs. This procedure is similar to the CGA procedure with the exception that experts are actually asked to generate the conceptual graph (as opposed to the ID generating the graph and merely presenting it to experts). This method may enable experts to recognize gaps in the knowledge and skills they provided and correct inadequacies.

After SMEs complete the Network module, data are stored on a floppy diskette and returned to the ID. The ID reviews the CE record and conceptual graphs for any glaring omissions in content. If any omissions are present, the ID can ask the expert to expand on any inadequacies.

The fifth module (i.e., Assess module) is used to validate the CE record and conceptual graph generated by SMEs. This is accomplished by having other experts in a domain review conceptual graphs generated by other experts. Multiple experts will be used to edit initial conceptual graphs as a method of modifying and validating the externalized knowledge structures.

The sixth module (i.e., Implicit Decompose module) is used to elicit any implicit knowledge experts have not previously described. This is accomplished by having additional SMEs review common and novel problems generated in the Explicit Decompose module. Experts are instructed to complete these tasks (either mentally or physically) and report their actions as well as their reasoning for each action. In essence, this is similar to a think-aloud concurrent observation technique; however, a knowledge engineer does not need to be present to collect data (i.e., DNA becomes the knowledge engineer).

The modules are repeated until IDs are satisfied with the content of the revised CE records and conceptual graphs. Efforts have been made to elicit SK, PS, and CK, as well as explicit and implicit forms of knowledge; however, the procedures outlined above must be empirically validated.

Future research directions for automated cognitive task procedures

DNA can be used as a tool to explore numerous basic and applied issues in cognitive psychology. Many of which will be explored during continued collaboration with Dr. Shute.

An important applied issue which must be addressed is whether DNA can obtain knowledge structures comparable to existing CTA methods in a shorter period of time. Furthermore, can knowledge structures obtained by DNA be easily translated into curricula? Research in this area must be conducted to determine the efficacy of DNA compared to other methods of CTA.

Can DNA obtain a comprehensive knowledge structure? The questioning dialog DNA utilizes is designed to elicit explicit forms of knowledge such as SK and CK (i.e., definitions of terms and relationships among terms), as well as implicit knowledge structures (i.e., PS elements that experts cannot easily recall). It is important to validate whether implicit knowledge can be elicited with retrospective observation of novel and common situations.

What is an expert? How do knowledge structures differ among various experts? Why do knowledge structures vary across experts? Research questions such as these are cognitive psychology issues which address basic issues. Several experts may be proficient in any given area, but research has shown that experts' knowledge structures vary greatly (Sanchez & Fraser, 1992).

Conclusion

Existing CTA methodologies are inefficient and often provide incomplete knowledge of the domain. The proposed automated cognitive task analysis procedure provides a principled approach to CTA which is based on theory and research. DNA is designed to elicit a comprehensive knowledge structure (i.e., both explicit and implicit versions of SK, PS, and CK) which can easily be translated into a curricula. However, extensive research must be conducted to assess the effectiveness and efficiency of such a method.

References

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. Psychological Review, 94, 192-210.

Anderson, J. R. (1988). The expert module. In M. C. Polson and J. J. Richardson (Eds.), Foundations of intelligent tutoring systems, (pp. 21-50). Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1990). Cognitive psychology and its implications (3rd ed.). New York, NY: Freeman.

Bloom, B. S. (1984). The 2-sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. Educational Researcher, 13, 4-16.

Burns, H. L., & Capps, C. G. (1988). Foundations of intelligent tutoring systems: An introduction. In M. C. Polson and J. J. Richardson (Eds.), Foundations of intelligent tutoring systems, (pp. 1-18). Hillsdale, NJ: Erlbaum.

Collins, H. M. (1985). Changing order: Replication and induction in scientific practice. London: Sage.

Durkin, J. (1994). Knowledge acquisition. In J. Durkin (Ed.), Expert systems: Design and development, (pp. 518-599). New York, NY: Macmillan.

Evans, J. (1988). The knowledge elicitation problem: A psychological perspective. Behavior and Information Technology, 7, 111-130.

Gamble, A., & Page, C. V. (1980). The use of artificial intelligence techniques in computer-assisted instruction: an overview. International Journal of Man-Machine Studies, 12, 259-282.

Gordon, S. E., Schmierer, K. A., & Gill, R. T. (1993). Conceptual graph analysis: Knowledge acquisition for instructional system design. Human Factors, 35, 459-481.

Graesser, A. C., & Clark, L. C. (1985). Structures and procedures of implicit knowledge. Norwood, NJ: Ablex.

Hall, E. P., Gott, S. P., & Pokorny, R. A. (1990). A procedural guide to cognitive task analysis: The PARI methodology. (Report No. AL/HR-TR-1990-0000). Air Force Materiel Command, Brooks Air Force Base, TX.

Hall, E. P., Gott, S. P., & Pokorny, R. A. (1995). A procedural guide to cognitive task analysis: The PARI methodology. (Report No. AL/HR-TR-1995-0108). Air Force Materiel Command, Brooks Air Force Base, TX.

Hayes-Roth, F., Waterman, D. A., & Lenat, D.B. (1983). Building expert systems. Reading, MA: Addison Wesley.

Kyllonen, P. C., & Christal, R. E. (1989). Cognitive modeling of learning abilities: A status report of LAMP. In R. Dillon & J. W. Pelligrino (Eds.), Testing: Theoretical and applied issues. San Francisco, CA: Freeman.

Kyllonen, P. C., & Shute, V. J. (1989). A taxonomy of learning skills. In P. L. Ackerman, R. J. Sternberg, & R. Glaser (Eds.), Learning and individual differences (pp. 117-163). New York, NY: Freeman.

McGraw, K. L., & Harbison-Briggs, K. (1989). Knowledge acquisition: Principles and guidelines. Englewood Cliffs, NJ: Prentice-Hall.

Moore, J. L., & Gordon, S. E. (1988). Conceptual graphs as instructional tools. Proceedings of the Human Factors Society 32nd Annual Meeting, 1289-1293.

Redding, R. E. (1989). Perspectives on cognitive task analysis: The state of the state of the art. Proceedings of the Human Factors Society 33rd Annual Meeting, 1348-1352.

Regian, J. W., & Shute, V. J. (1992). Automated instruction as an approach to individualization. In J. W. Regian and V. J. Shute (Eds.), Cognitive approaches to automated instruction, (pp. 1-13). Hillsdale, NJ: Erlbaum.

Sanchez, J. I., & Fraser, S. L. (1992). On the choice of scales for task analysis. Journal of Applied Psychology, 77, 545-553.

Shute, V. J. (1995). SMART: Student modeling approach for responsive tutoring. User Modeling and User-Adapted Interaction, 5, 1-44.

Shute, V. J., Torreano, L. A. (1995). Cognitive task analysis made simple: Decompose, network, assess (DNA). Unpublished manuscript.

VanLehn, K. (1988). Student modeling. In M. C. Polson and J. J. Richardson (Eds.), Foundations of intelligent tutoring systems, (pp. 55-76). Hillsdale, NJ: Erlbaum.

Figure 1. Learning hierarchy of Blackjack constructed by Shute, Sugrue, and Willis.

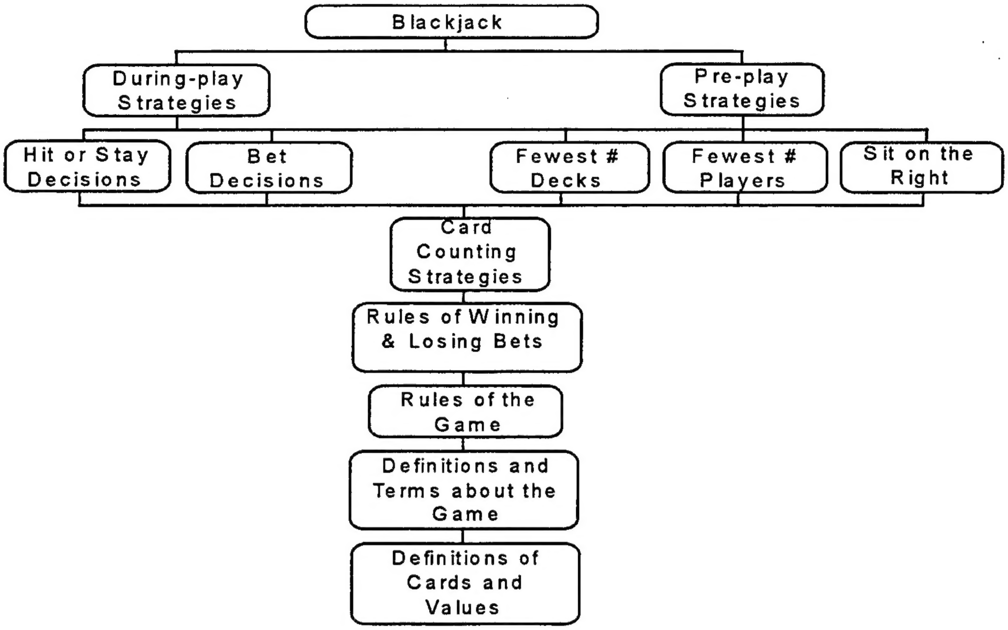


Figure 2. Conceptual graph of searching a videotape for specific information adapted from Gordon, Schmierer, and Gill (1993).